I.M. YAGLOM

NOUVELLES ORIENTATIONS DES MATHÉMATIQUES

I. Yaglom, B. Trakhtenbrot,
E. Ventsel, A. Solodovnikov

NOUVELLES ORIENTATIONS DES MATHÉMATIQUES

Editions Mir · Moscou

AVANT-PROPOS

Le présent ouvrage constitue le sixième opuscule de la collection « Initiation aux mathématiques » publiée par les Editions Mir. Cette collection est destinée à un large auditoire et surtout aux élèves du secondaire. Les fascicules faisant partie de cet ouvrage ne sont pas liés entre eux et peuvent donc être lus dans n'importe quel ordre.

Le présent opuscule réunit les articles où l'on traite les nouvelles applications des mathématiques, qui ont vu ces dernières années un large essor.

Dans le premier fascicule « Algèbre non ordinaire » le professeur I. Yaglom traite l'algèbre dite de Boole, qui constitue la base de la logique mathématique. La portée de la logique mathématique, qui a fait son apparition au XIXe siècle, s'est considérablement accrue avec la création des machines à calculer électroniques: les tentatives de charger les machines à calculer électroniques de certaines fonctions auparavant considérées comme le privilège de l'homme (aujourd'hui les machines électroniques traduisent des textes, jouent aux dames et aux échecs, et même démontrent des théorèmes!) nous ont nécessités à faire une description formelle du processus même de raisonnement (de déduction logique). Ce fascicule du présent ouvrage est le plus élémentaire.

Dans le second fascicule « Algorithmes et résolution de problèmes par des machines », dû au professeur B. Trakhtenbrot, on examine le problème de la formalisation du processus de la pensée et on étudie les principes des machines à calculer électroniques. Bien que ce fascicule ne dépende pas du premier, les idées exposées dans celui-ci concernant l'algèbre des propositions aideront le lecteur à mieux assimiler les notions d'algorithmes et le principe de fonctionnement des machines à calculer électroniques.

Dans le troisième fascicule « Eléments de la théorie des jeux » le professeur E. Ventsel expose les fondements de la

théorie des jeux, branche toute récente des mathématiques due au remarquable mathématicien J. von Neumann, qui a des applications importantes. Dans le § 5 on examine les rapports entre la théorie des jeux et les méthodes de la programmation linéaire servant à appliquer les mathématiques à la résolution des problèmes économiques.

Le quatrième fascicule « Systèmes d'inégalités linéaires » de A. Solodovnikov traite les fondements de la programmation linéaire.

Il est à remarquer que les fascicules 3-4, destinés aux lecteurs ayant une formation mathématique assez solide et désireux de se faire une idée précise de l'appareil mathématique pratique, sont plus compliqués que les fascicules 1-2.

I. Yaglom

ALGÈBRE NON ORDINAIRE

§ 1. ALGÈBRE DES NOMBRES ET ALGÈBRE DES ENSEMBLES

Aux leçons d'arithmétique et d'algèbre, on étudie les nombres de toute nature. Les élèves des classes élémentaires, en apprenant les nombres entiers, n'éprouvent pratiquement aucune difficulté sérieuse : la plupart d'entre eux en possèdent déjà certaines notions. Mais voici qu'apparaissent des « nombres » toujours nouveaux ; nous nous y sommes déjà habitués, et ils ne nous étonnent plus, bien qu'à chaque étape de l'extension de la notion de nombre nous disions adieu à des illusions si chères à notre cœur. Un nombre (entier) montre combien d'objets contient tel ou tel ensemble ; par exemple, combien y a-t-il de pommes dans un panier ? de pages dans un livre ? ou d'élèves dans une classe ? Mais qu'en est-il des fractions ? Il va de soi qu'une classe ne peut contenir $33\,{}^{1}/_{3}$ élèves de même qu'il ne peut y avoir $3^{1}/_{4}$ assiettes sur la table. Bien sûr que non. Mais sur la table on peut bien poser $4\,{}^{1}/_{2}$ pommes, un film peut bien durer $1\,{}^{3}/_{4}$ heure, et une bibliothèque peut bien renfermer $6\,{}^{1}/_{2}$ livres (certes, ceci en dit long sur les soins apportés à cette bibliothèque, mais n'est nullement en contradiction avec le bon sens !). A peine nous sommes-nous habitués à ce que des objets quelconques peuvent exister en quantité fractionnaire, qu'apparaissent des nombres négatifs. Cependant, on ne peut dire qu'une bibliothèque contient — 3 livres, ceci aurait été contraire à la nature. Mais, en revanche, un thermomètre peut bien indiquer —5°, ou quelqu un n'a, par exemple, que 5 francs ce qui est triste... pour lui, mais non pas pour les mathématiques. Et voici que dans les classes terminales on se heurte, enfin, à des nombres tout à fait « affreux » : d'abord les nombres irrationnels tel $\sqrt{2}$ par exemple, puis les nombres imaginaires tel $1 + 2i$ *) par exemple ; le

*) **Aujourd'hui les nombres tels que $1 + 2i$ sont souvent appelés nombres *complexes*, le terme *imaginaire* (ou *imagi-***

qualificatif même de ces nombres est significatif de la disposition des hommes à leur égard, du moins tant qu'ils ne s'y sont pas habitués. Si le lecteur ne connaît encore pas ces nombres, et s'il ne commence qu'à s'en faire une idée *) cela ne l'empêche aucunement d'assimiler le contenu de ce livre. Bien que les nombres irrationnels et imaginaires soient fort éloignés de l'idée primitive du nombre en tant que caractéristique d'une collection d'objets quelconques, on les appelle tout de même aussi « nombres ».

Mais qu'ont-ils donc de commun ces nombres pour être tous groupés sous la même appellation? Le seul trait commun est qu'on peut leur appliquer les opérations d'addition et de multiplication **). Cette ressemblance n'est cependant qu'apparente: bien que nous puissions additionner et multiplier les nombres de toute nature, ces opérations n'ont pas toujours le même sens. Ainsi, additionner deux nombres entiers positifs a et b c'est déterminer le nombre d'objets dans la réunion de deux ensembles dont le premier contient a et le deuxième b objets: si, par exemple, dans une classe il y a 35 élèves et dans une autre 39, alors dans ces deux classes il y a $35 + 39 = 74$ élèves (voir fig. 1). D'une façon analogue, multiplier deux nombres entiers positifs a et b c'est déterminer le nombre d'objets dans la réunion de a ensembles formés de b objets chacun: ainsi, par exemple, si dans une école il y a 3 premières classes de 36 élèves chacune, cela signifie

---

naire pur) étant réservé aux nombres tels que $2i$ ou $-\sqrt{2}i$ (par opposition, on appelle *réels* les nombres de la forme 1, $-3/2$ ou $\sqrt{2}$).

*) Un très bon ouvrage est entièrement consacré à des nombres de différentes espèces. A. Niven, *Nombres rationnels et irrationnels*, Editions « Mir », 1966.

**) Mais non pas celles de soustraction ou de division: si nous ne connaissons que des nombres positifs, nous ne pourrons pas soustraire 5 de 3, et, d'une façon analogue, si nous ne connaissons que des nombres entiers, nous ne pourrons pas diviser 7 par 4.

que dans cette école il y a $3 \cdot 36 = 108$ élèves de première en tout (voir fig. 2). Néanmoins, les définitions de l'addition et de la multiplication énoncées ainsi ne permettent pas d'étendre ces opérations aux opérations sur les fractions ou à celles sur les nombres négatifs (ni *a fortiori* sur les nombres irrationnels et imaginaires).

Ainsi, nous sommes amenés à la conclusion suivante: les nombres de toute nature sont appelés « nombres » du fait qu'on peut les additionner et multiplier, mais que les opérations même d'addition et de multiplication pour les différentes espèces de nombres s'effectuent d'une façon tout à fait différente. Il est quand même à souligner que cette conclusion est trop hâtive: en effet l'addition des nombres entiers et l'addition des fractions ne diffèrent pas autant. Plus précisément, si les *définitions* de ces opérations sont différentes, leurs *propriétés* n'en sont pas moins absolument identiques. Ainsi, *quelle que soit* la nature d'un nombre, on a

$$a + b = b + a \quad \text{et} \quad ab = ba\,;$$

loi commutative pour l'addition — loi commutative pour la multiplication

$$(a + b) + c = a + (b + c) \quad \text{et} \quad (ab)\,c = a\,(bc)\,;$$

loi associative pour l'addition — loi associative pour la multiplication

il existe dans tous les cas deux nombres « particuliers » 0 et 1, tels que

$$a + 0 = a \quad \text{et} \quad a \cdot 1 = a$$

pour tout a. Donc, on peut en conclure que l'objet de l'algèbre moderne est d'étudier certains systèmes (différents) de nombres pour lesquels sont définies les opérations d'addition et de multiplication vérifiant les lois ci-dessus et certaines autres comme, par exemple:

$$(a + b)\,c = ac + bc,$$

loi distributive pour la multiplication par rapport à l'addition

où a, b, c sont des nombres de *n'importe quelle* nature;

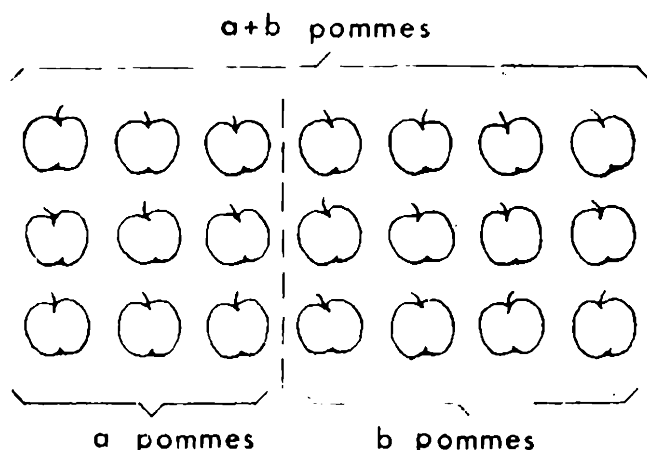


Fig. 1

La présence dans les systèmes de nombres de deux opérations, l'addition et la multiplication, crée un certain parallélisme, qui est d'autant plus prononcé que les propriétés de l'addition rappellent sous beaucoup de rapports celles de la multiplication. Ce parallélisme se manifeste, par exemple, dans cette proportion « curieuse » dans laquelle chacun

$$\frac{\text{addition}}{\text{soustraction}} = \frac{\text{multiplication}}{?}$$

voudrait remplacer le point d'interrogation par « division » sans même réfléchir ce que signifie cette proportion. Ce même parallélisme se traduit également par le fait que les élèves, et même parfois leurs parents confondent les termes « nombre opposé » (c'est-à-dire le nombre $-a$ dont la somme avec a donne 0) et « nombre inverse » (c'est-à-dire le nombre $1/a$ dont le produit par a donne 1). Ce parallélisme apparaît, enfin, dans la similitude des propriétés de la progression arithmétique (qui est une suite de nombres tels que la différence entre l'un d'eux et son antécédent est une quantité constante) et celles de la progression géométrique (qui est une suite de nombres dans laquelle le quotient de deux termes qui se suivent est une quantité constante).

Néanmoins, ce parallélisme n'existe pas toujours. Ainsi, par exemple, le nombre 0 joue un rôle particulier tant dans

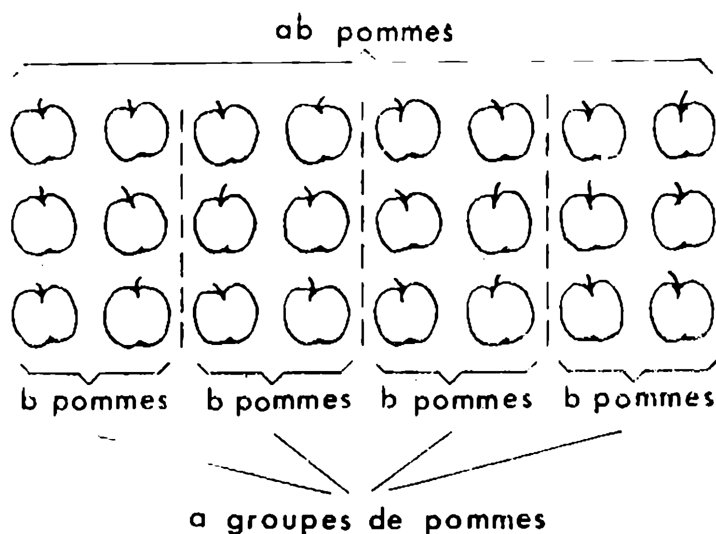


Fig. 2

l'addition que dans la multiplication: en effet pour tout nombre a

$$a \cdot 0 = 0$$

(d'où il s'ensuit, entre autres, qu'on ne peut pas diviser un nombre non nul par 0). Cependant si l'on remplace dans cette dernière égalité la multiplication par l'addition, et 0 par 1, on aura l' « égalité » absurde suivante

$$a + 1 = 1$$

qui n'est valable que pour $a = 0$ *). D'autre part, si l'on remplace dans la loi distributive $(a + b)\,c = ac + bc$ l'addition par la multiplication et vice versa, on obtient l'« égalité »

$$ab + c = (a + c)\,(b + c)$$

qui n'a pas de sens dans le cas général. [Etant donné que

$$(a + c)\,(b + c) = ab + ac + bc + c^2 =$$
$$= ab + c\,(a + b + c),$$

*) Si l'égalité $a + 1 = 1$ avait lieu pour tout a, il serait impossible de retrancher 1 d'un nombre quelconque différent de 1. En réalité cela n'est certainement pas vrai vu que, par exemple, $3 - 1 = 2$.

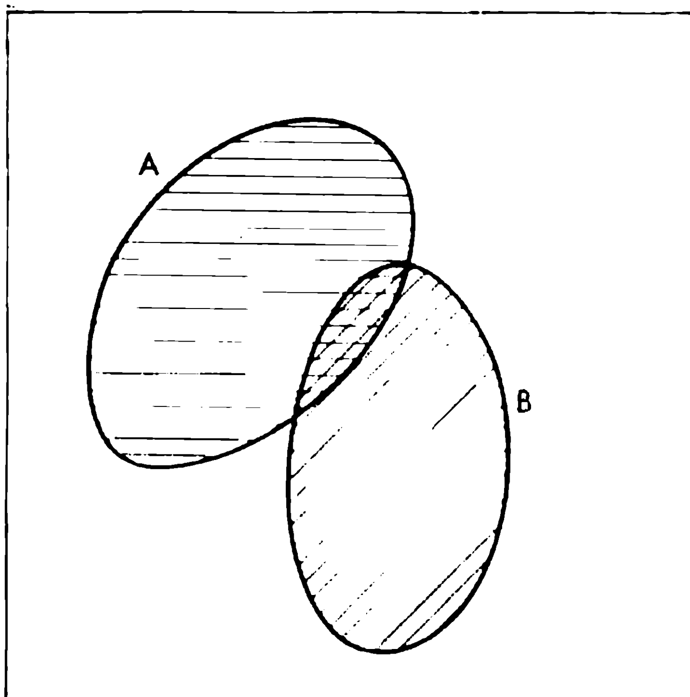


Fig. 3

l'égalité $(a + c)(b + c) = ab + c$ n'est vraie que si $c = 0$ ou si $a + b + c = 1$.]

Mais il existe en algèbre d'autres systèmes, outre les systèmes numériques, dans lesquels les opérations d'addition et de multiplication ont plus de commun entre elles que les mêmes opérations sur les nombres. Telle est, par exemple, l'« *algèbre des ensembles* », une branche fort importante de l'algèbre moderne. Par *ensemble* on entend un assemblage d'objets quelconques, appelés *éléments d'un ensemble*: on peut dire l'« ensemble des élèves d'une classe », l'« ensemble des points d'un cercle », l'« ensemble des points d'un carré », l'« ensemble des éléments de la classification périodique de Mendéléev », l'« ensemble des nombres pairs », l'« ensemble des notes dans un journal de classe », l'« ensemble des éléphants en Inde », l'« ensemble des fautes de grammaire dans une composition », etc. Il est aisé de comprendre comment on peut définir l'addition de deux ensembles: *par somme* $A + B$ *des ensembles* A *et* B *on entend tout simplement la réunion de ces deux ensembles*. Si A, par exemple, est l'en-

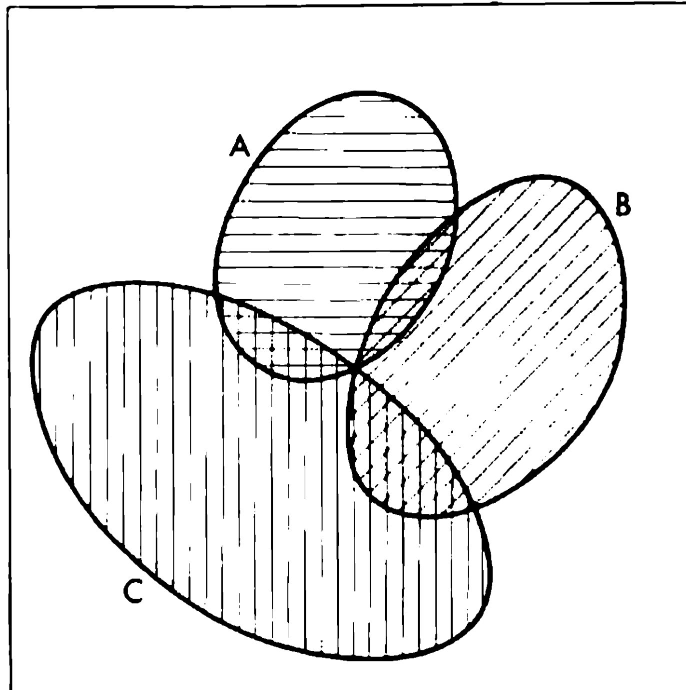


Fig. 4

semble des garçons d'une classe et B celui des filles de cette classe, $A + B$ sera alors l'ensemble des élèves de cette classe ; si A est l'ensemble de tous les nombres entiers positifs pairs et B l'ensemble des nombres multiples de 3, l'ensemble $A + B$ s'écrit

$$\{2, 3, 4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22 \ldots\}$$

si, enfin, l'ensemble A se compose des points de la figure hachurée horizontalement (voir fig. 3) et l'ensemble B des points de la figure hachurée obliquement, l'ensemble $A + B$ est constitué par toute la figure hachurée (fig. 3). De la figure 3, il résulte que pour chaque couple d'ensembles A et B a lieu la relation suivante

$$A + B = B + A,$$

c'est-à-dire que l'*addition des ensembles est régie par une loi commutative*. Par ailleurs, quels que soient les ensembles A, B et C, on a toujours

$$(A + B) + C = A + (B + C),$$

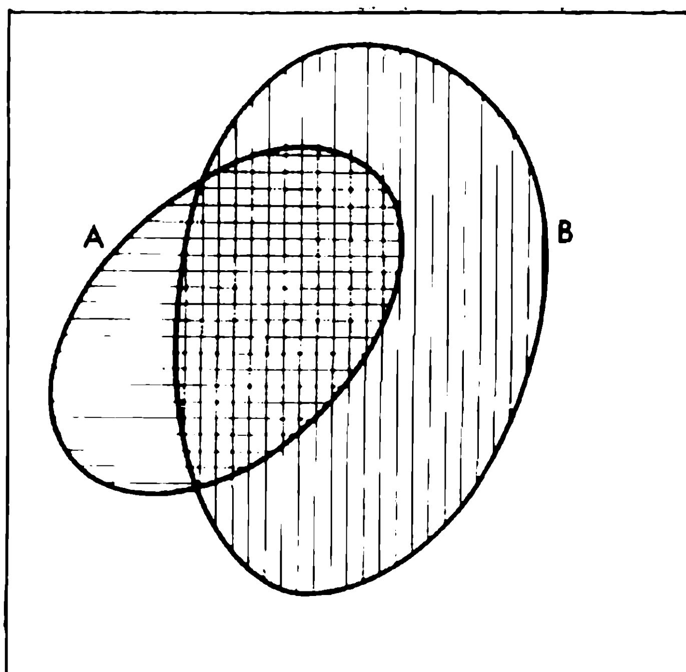


Fig. 5

c'est-à-dire que l'*addition des ensembles est une opération associative*. L'ensemble $(A + B) + C$ (ou $A + (B + C)$) peut tout simplement être écrit $A + B + C$ sans parenthèses : il représente donc la réunion des trois ensembles A, B et C (l'ensemble $A + B + C$ représenté sur la figure 4 se confond avec toute la figure hachurée).

Convenons d'appeler *produit AB des ensembles A et B la partie commune ou l'intersection de ces ensembles*. Si, par exemple, A est l'ensemble des joueurs d'échecs d'une classe quelconque et B l'ensemble des nageurs de cette classe, alors AB est l'ensemble des joueurs d'échecs sachant nager ; si A est l'ensemble des nombres entiers positifs pairs et B l'ensemble des multiples de 3, l'ensemble AB sera composé des multiples de 6 :

$$\{6, 12, 18, 24 \ldots\};$$

si, enfin, l'ensemble A est constitué par les points de la figure hachurée horizontalement sur la figure 5, l'ensemble B, par les points de la figure hachurée verticalement, l'ensemble

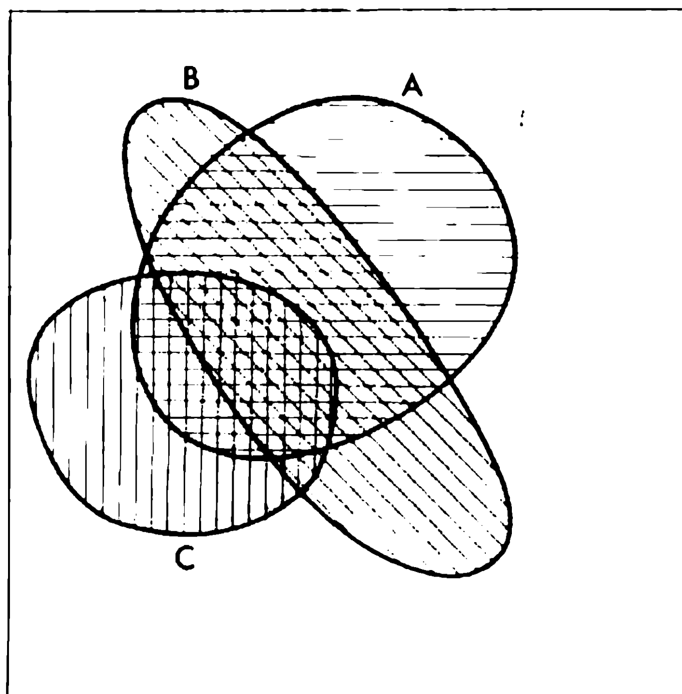


Fig. 6

AB sera constitué par deux sortes de hachures. Il est clair que le *produit des ensembles vérifie la loi commutative*, c'est-à-dire quels que soient les ensembles A et B

$$AB = BA,$$

(cf. fig. 5 ; il est aisé de voir que l'« ensemble AB des joueurs d'échecs sachant nager » et l'« ensemble BA des nageurs sachant jouer aux échecs » sont le même ensemble). Ensuite, il est non moins évident que le *produit des ensembles vérifie la loi associative*, c'est-à-dire pour les trois ensembles A, B et C

$$(AB)\,C = A\,(BC).$$

On peut donc écrire indifféremment $(AB)\,C$, $A\,(BC)$ ou ABC ; l'ensemble produit représente la partie commune ou l'intersection des trois ensembles A, B et C (l'ensemble

ABC est marqué par des hachures horizontales, verticales et obliques sur la figure 6*)).

Il est à remarquer qu'*entre trois ensembles quelconques* A, B *et* C *a lieu* la relation suivante

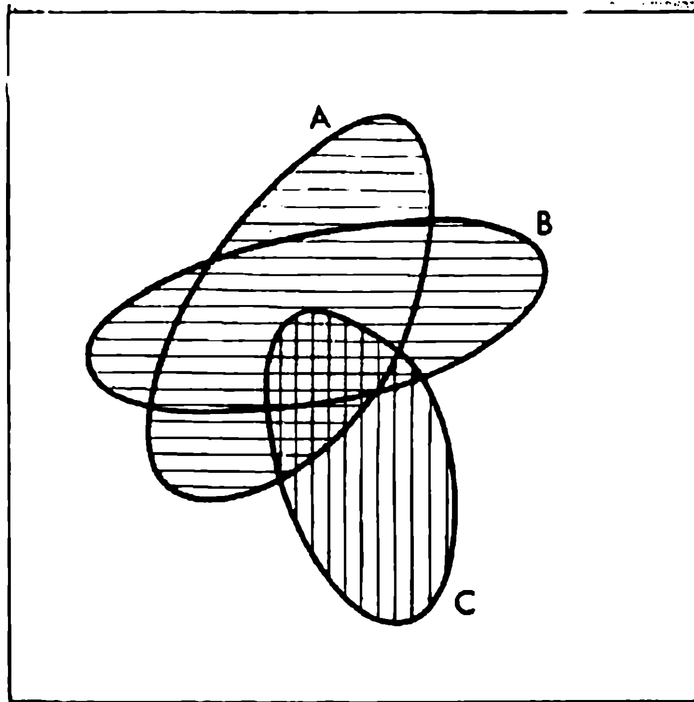
$$(A + B)\, C = AC + BC,$$

c'est-à-dire a lieu la *loi distributive*. En effet, si, par exemple, dans une classe A est l'ensemble des joueurs d'échecs, B celui des joueurs de dames et C celui des nageurs, l'ensemble $A + B$ est alors la réunion de deux ensembles : celui des joueurs d'échecs et celui des joueurs de dames, c'est-à-dire l'ensemble des élèves sachant jouer soit aux échecs soit aux dames (ou aux échecs et aux dames à la fois). L'ensemble $(A + B)\, C$ s'obtient à partir de l'ensemble $A + B$, si l'on réunit les élèves qui savent encore nager. Nous obtiendrons le même ensemble si nous composons l'ensemble $AC + BC$ en réunissant l'ensemble AC des joueurs d'échecs sachant nager avec l'ensemble BC des joueurs de dames sachant nager.

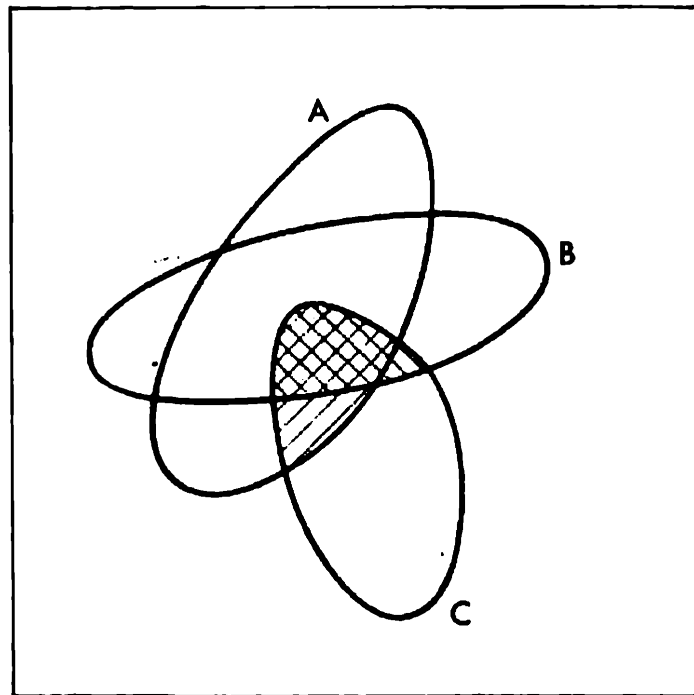
Il est probable qu'une telle interprétation de la loi distributive puisse sembler un peu compliquée. Aussi aurons-nous recours à une illustration graphique. Sur la figure 7,*a* l'ensemble $A + B$ est hachuré horizontalement, l'ensemble C verticalement de sorte que l'ensemble $(A + B)\, C$ est quadrillé. Les ensembles représentés sur la figure 7,*b* sont marqués de hachures inclinées respectivement à droite et à gauche ; l'ensemble $AC + BC$ se confond alors avec toute

---

*) Voici encore un exemple illustrant que le produit des ensembles vérifie la loi associative. Soient A l'ensemble des multiples de 2, B celui des multiples de 3 et C celui des multiples de 5. AB est alors l'ensemble des multiples de 6, $(AB)\, C$ celui des multiples de 6 et de 5, c'est-à-dire de 30. D'autre part, BC est l'ensemble des multiples de 15 et $A\, (BC)$ celui des nombres pairs multiples de 15, autrement dit, $A\, (BC)$ est de nouveau l'ensemble des multiples de 30.



a)



b)

Fig. 7

la figure hachurée. On constate aisément que l'ensemble $AC + BC$ de la figure 7,*b* ne se distingue en rien de l'ensemble $(A + B)\ C$ marqué d'une hachure double sur la figure 7,*a*.

On comprend aisément quel « ensemble » est appelé à jouer le rôle du zéro. L'addition de l'ensemble O (que nous désignerons par la lettre O) ne doit modifier aucun ensemble, donc, il ne doit renfermer *aucun* élément, c'est-à-dire il est « vide ». On sera sans doute tenté d'éliminer l'*ensemble vide* de tout examen en raisonnant comme suit : si l'ensemble O ne possède aucun élément, ce n'est plus un ensemble du tout, alors à quoi bon l'étudier ! Or, en agissant de la sorte rien ne nous empêchera d'exclure par la suite 0 de l'ensemble des nombres : car une « collection » de zéro objets est également « vide », et il paraît absurde de parler du « nombre » d'objets qu'il contient. Mais en réalité il y a toutes les raisons d'en parler. S'il n'y avait pas de zéro on ne pourrait jamais retrancher un nombre d'un autre (en effet dans ce cas la différence 3 — 3 ne serait égale à rien) ; comment noterait-on, par exemple, le nombre 108 dans le système numérique décimal (une centaine, huit unités et aucune dizaine !) ; de tels exemples ne manquent pas, et il n'est donc pas surprenant que l'introduction du zéro soit l'un des événements les plus marquants dans l'histoire des mathématiques. D'une façon analogue, si l'on considère que l'ensemble vide O n'est pas un ensemble, on ne pourra jamais définir le produit (ou intersection) de deux ensembles *quelconques* ; ainsi, par exemple, l'intersection des ensembles A et B, représentés sur la figure 8, est vide, de même qu'est vide l'intersection de l'ensemble des meilleurs élèves d'une classe quelconque avec, disons, celui des éléphants. Et d'ailleurs, si l'on renonce à la notion d'« ensemble vide », comment pourra-t-on parler des ensembles en général : et si, par exemple, l'« ensemble des élèves portant le nom d'André d'une classe quelconque » s'avère vide, c'est-à-dire qu'un tel ensemble n'existe pas du tout ?

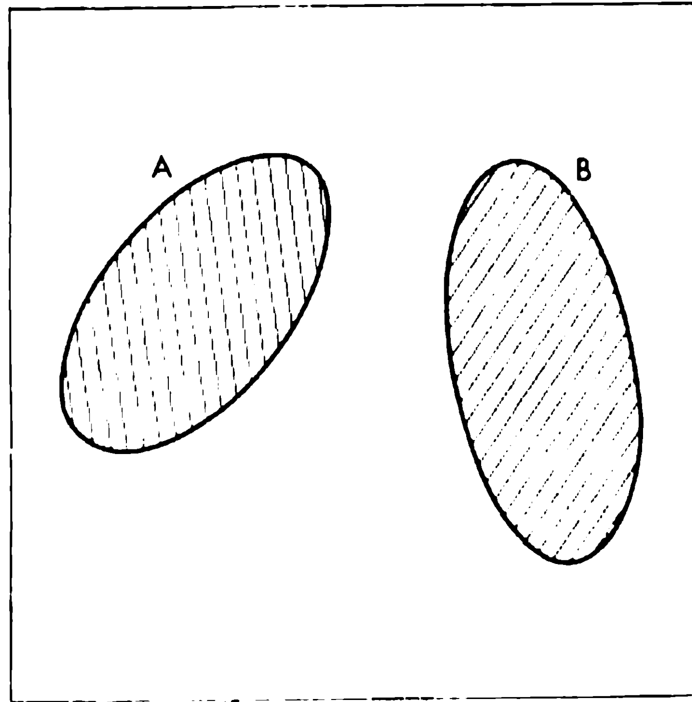


Fig 8.

Il est clair que si O est un ensemble vide, alors pour tout ensemble A

$$A + O = A.$$

Il est non moins clair que, quel que soit l'ensemble A, on a toujours

$$AO = O$$

puisque l'intersection de tout ensemble A et de l'ensemble O (disons, l'intersection de l'ensemble des filles d'une classe quelconque et de celui des élèves dont la taille dépasse 2,5 m) est nécessairement vide!

Pour l'« ensemble unité », les choses sont un peu plus compliquées. Cet ensemble (que nous allons désigner par la lettre I) doit être tel que le produit (c'est-à-dire l'intersection) de cet ensemble et de tout ensemble A doive se confondre avec A. Il en résulte que notre ensemble I doit contenir tous les éléments en général de tous les ensembles A ! Il est évident qu'un tel ensemble ne peut exister que dans le cas où l'on se limite aux ensembles A dont les éléments sont

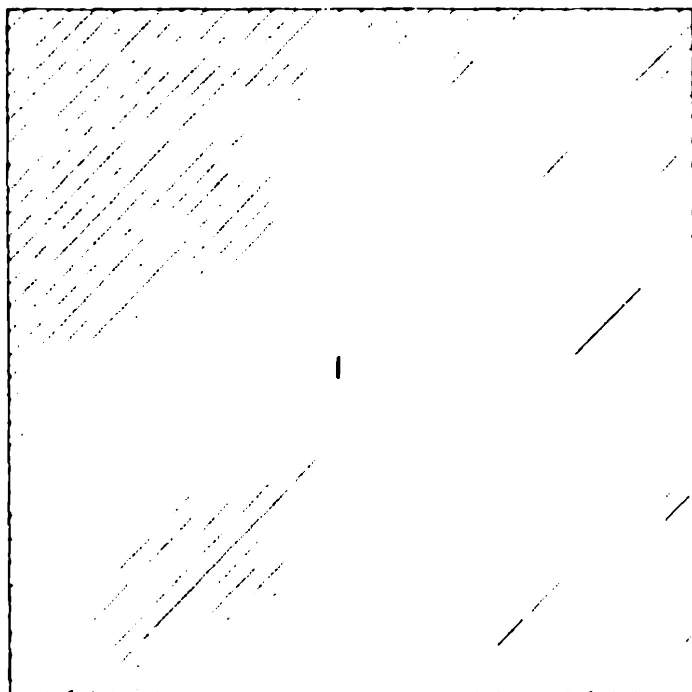


Fig. 9

puisés à une réserve déterminée quelconque d'« objets »: les ensembles des élèves d'une école ou d'une classe (*A* peut désigner l'ensemble des meilleurs élèves, *B* l'ensemble des joueurs d'échecs) ; les ensembles des nombres entiers positifs (*A* étant, par exemple, l'ensemble des nombres pairs, *B* celui des nombres premiers) ; les ensembles des points qui forment dans un carré quelconque des figures pareilles à celles représentées sur les figures 3-8. Ceci dit, nous entendrons par l'« ensemble unité » *I* « le plus grand ensemble » renfermant *tous* les « objets » considérés: l'ensemble de *tous* les élèves d'une école ou d'une classe, l'ensemble de *tous* les nombres entiers positifs ou, enfin, l'ensemble de *tous* les points d'un carré (fig. 9). Dans l'« algèbre des ensembles » l'ensemble *I* est dit *ensemble unité* ou *ensemble universel.* Il est évident que pour tout ensemble *A* « plus petit » (et même pour celui qui coïncide avec *I*) on aura l'égalité

$$AI = I,$$

conformément à la condition qui définit l'unité.

Nous sommes amenés donc à conclure que les opérations sur les ensembles que nous avons définis rappellent sous beaucoup de rapports celles de l'algèbre des nombres sans pour autant les doubler entièrement. En effet, dans l'algèbre des ensembles ont lieu presque toutes les lois fondamentales des nombres, mais à côté de ces lois il en existe d'autres qui peuvent étonner de prime abord. Nous avons déjà noté que la règle qui résulte de l'égalité $a \cdot 0 = 0$ si l'on remplace la multiplication par l'addition et 0 par 1, n'est pas valable pour les nombres: pour tous les nombres au moins a on a $a + 1 \neq 1$. Par contre, dans l'algèbre des ensembles il en va autrement: l'égalité

$$A + I = I$$

est, elle, toujours vérifiée. En effet, l'ensemble I étant par définition « le plus grand », il est donc impossible de l'augmenter davantage: quel que soit l'ensemble A (parmi ceux considérés) qu'on ajoute à l'ensemble unité I on aura toujours le même ensemble I.

En remplaçant ensuite dans la relation $(a + b)\, c = ac + bc$ l'addition par la multiplication et vice versa on obtient une « égalité » absurde $ab + c = (a + c)\,(b + c)$ qui est presque toujours incorrecte pour les nombres. Il en est autrement dans l'algèbre des ensembles: ici a toujours lieu l'égalité

$$AB + C = (A + C)\,(B + C),$$

qui traduit la *deuxième loi distributive* (valable pour l'addition par rapport à la multiplication) (quels que soient les ensembles A, B et C). En effet, soient de nouveau A l'ensemble des joueurs d'échecs, B celui des joueurs de dames et C celui des nageurs. Dans ce cas l'intersection AB des ensembles A et B se compose de tous les élèves sachant jouer aux échecs et aux dames et la réunion $AB + C$ des ensembles AB et C se compose des élèves sachant jouer aux échecs et

(affectés du signe « + » ou « — »). Il n'est donc pas nécessaire de réapprendre toutes ces lois, mais seulement de les compléter compte tenu de l'élargissement du domaine des nombres étudiés. Par contre, lorsqu'on passe des nombres aux ensembles la situation change: certaines lois de l'algèbre des ensembles ne sont pas valables pour les nombres *).

---

*) C'est cette différence justement entre les lois de l'algèbre des ensembles et les lois des nombres qui est à l'origine de l'utilisation des symboles $\cup$ et $\cap$ pour désigner respectivement l'addition et la multiplication des ensembles (c'est-à-dire leur réunion et leur intersection), au lieu des signes « + » et « · »: ainsi $A \cup B$ est la réunion des ensembles A et B, $A \cap B$, leur intersection. Etant donné que dans ce qui suit nous aborderons également d'autres systèmes algébriques où l'« addition » et la « multiplication » obéissent aux mêmes lois que dans l'algèbre des ensembles, nous abandonnerons les symboles $\cup$ et $\cap$, qui sont exclusivement utilisés en théorie des ensembles. Ainsi, pour souligner l'affinité des algèbres examinées avec celle que nous avons apprise au lycée nous utiliserons par la suite les signes ordinaires d'addition et de multiplication. Il semble toutefois utile de donner ici les lois fondamentales de l'algèbre des ensembles sous la forme généralement admise en théorie des ensembles:

$$A \cup B = B \cup A \quad \text{et} \quad A \cap B = B \cap A:$$

lois commutatives

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{et} \quad (A \cap B) \cap C = A \cap (B \cap C);$$

lois associatives

$$A \cup O = A \quad \text{et} \quad A \cap I = A,$$
$$A \cup I = I \quad \text{et} \quad A \cap O = O;$$

propriétés de l'ensemble vide O et de l'ensemble unité I

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad \text{et} \quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C);$$

lois distributives

$$A \cup A = A \quad \text{et} \quad A \cap A = A.$$

lois idempotentes

Ecrivons ces nouvelles lois. On rapporte à ces lois la relation suivante

$$A + I = I$$

qui marque une différence profonde existant entre l'ensemble unité I et le nombre 1. La deuxième loi distributive de l'algèbre des ensembles nous permet de « chasser les parenthèses » de façon très particulière

$$(A + C)(B + C) = AB + C;$$

ainsi nous pouvons écrire, par exemple,

$$(A + D)(B + D)(C + D) = [(A + D)(B + D)](C + D) = \\ = (AB + D)(C + D) = (AB)C + D = ABC + D.$$

Et, enfin, les lois idempotentes sont toutes nouvelles pour le lecteur

$$A + A = A \text{ et } AA = A.$$

Ces lois autorisent parfois à affirmer que dans l'*algèbre des ensembles il n'existe ni exposants, ni coefficients*. En effet, on a

$$\underbrace{A + A + \ldots + A}_{n \text{ fois}} = A \quad \text{et} \quad \underbrace{A \cdot A \cdot \ldots \cdot A}_{n \text{ fois}} = A,$$

quels que soient A et n; pour cette raison,

$$(A + B)(B + C)(C + A) = \\ = ABC + AAB + ACC + AAC + BBC + ABB + \\ + BCC + ABC = (ABC + ABC) + (AB + AB) + \\ + (AC + AC) + (BC + BC) = ABC + AB + AC + BC$$

(cf. avec l'exercice 6 ci-dessous).

Exercices

Démontrer les égalités suivantes où les majuscules désignent des ensembles (la lettre O symbolise toujours l'ensemble vide, I l'ensemble unité):

1. $(A + B)(A + C)(B + D)(C + D) = AD + BC$.

2. $A(A+B)=A$.
3. $AB+A=A$.
4. $A(A+C)(B+C)=AB+AC$.
5. $A(A+I)(B+O)=AB$.
6. $(A+B)(B+C)(C+A)=AB+BC+CA$.
7. $(A+B)(B+C)(C+D)=AC+BC+BD$.
8. $(A+B)(A+I)+(A+B)(B+O)=A+B$.
9. $(A+B)(B+I)(A+O)=A$.
10. $(A+B+C)(B+C+D)(C+D+A)=AB+AD+$
$+BD+C$.

E x e p l e : $A(A+C)(B+C) \underset{\text{assoriativité de la miltiplication}}{=} A[(A+C)(B+C)]=$

$\underset{\text{2me loi distributive}}{=} A(AB+C) \underset{\text{commutativité dc la multiplication}}{=} (AB+C)A=(AB)A+CA \underset{\text{1re loi distributive}}{=}$

$\underset{\text{commutativité et associativité de la multiplication}}{=} (AA)B+AC \underset{\text{idempotence de la multiplication}}{=} AB+AC.$

§ 2. ALGÈBRES DE BOOLE

Groupons toutes les lois de l'algèbre des ensembles que nous connaissons:

$A+B=B+A$ et $AB=BA$;
lois commutatives

$(A+B)+C=A+(B+C)$ et $(AB)C=A(BC)$;
lois associatives

$(A+B)C=AC+BC$ et $AB+C=(A+C)(B+C)$;
lois distributives

$A+A=A$ et $AA=A$.
lois idempotentes

De plus, en algèbre des ensembles il y a deux éléments (ensembles) « particuliers », O et I tels que

$$A + O = A \quad \text{et} \quad AI = A;$$
$$A + I = I \quad \text{et} \quad AO = O.$$

Ces lois (opérations) sont proches des lois ordinaires de l'algèbre des nombres, mais elles ne coïncident pas avec ces dernières; l'algèbre des ensembles est aussi une algèbre, mais vis-à-vis de l'algèbre que nous avons connue à l'école, c'est une algèbre non ordinaire.

Notons que l'*algèbre* ordinaire des nombres ne constitue pas une seule mais plusieurs algèbres: il existe l'« algèbre des nombres entiers positifs », l'« algèbre des nombres rationnels (c'est-à-dire des nombres entiers et fractionnaires) », l'« algèbre des nombres relatifs (c'est-à-dire des nombres positifs et non positifs) »; il existe également l'« algèbre des nombres réels (c'est-à-dire des nombres rationnels et irrationnels) », l'« algèbre des nombres complexes (des nombres réels et imaginaires) », etc. Toutes ces « algèbres » diffèrent par les nombres sur lesquels on opère, et par les définitions de ces opérations (c'est-à-dire l'addition et la multiplication), les propriétés fondamentales de ces opérations étant cependant les mêmes dans tous les cas. Que peut-on dire de l'algèbre non ordinaire des ensembles? Possède-t-elle une forme unique ou bien se compose-t-elle également d'une série d'« algèbres » semblables qui diffèrent l'une de l'autre par les éléments qui interviennent dans les opérations de calcul et par la définition de ces opérations (que nous appelons toujours addition et multiplication)?

Le lecteur devine déjà, et non sans raison d'ailleurs, qu'il existe plusieurs algèbres semblables à celle des ensembles (c'est-à-dire régies par les mêmes règles que celles de l'algèbre des ensembles). Il est à souligner avant tout que les algèbres des ensembles elles-mêmes sont assez diverses: ainsi, par exemple, on peut parler de l'« algèbre des ensem-

bles des élèves d'une classe quelconque », de l'« algèbre des ensembles des bêtes du zoo de Moscou » (il s'agira bien sûr d'une tout autre algèbre !), de l'« algèbre des ensembles des nombres (quelconques) », de l'« algèbre des ensembles des points d'un carré » (voir fig. 3-10), de l'« algèbre des ensembles des étoiles dans le ciel », ou, enfin, de l'« algèbre des ensembles des livres dans une bibliothèque », etc. Mais il existe d'autres exemples d'algèbres aux propriétés identiques.

Avant de passer à ces exemples, attardons-nous sur quelques détails d'importance. Avant d'examiner les exemples qui vont suivre il importe de savoir que *définir dans un ensemble quelconque des objets* (*ou éléments*) a, b, . . . *les opérations d'addition et de multiplication signifie formuler les règles suivant lesquelles deux objets quelconques* a *et* b *sont confrontés avec deux autres objets* c *et* d *appelés respectivement somme et produit de* a *et* b :

$$c = a + b, \ d = ab.$$

Nous allons choisir ces règles de façon que soient vérifiées toutes les lois de l'algèbre des ensembles. Et de plus, nous ne devons pas poser la question : pourquoi la somme de a et b est-elle égale à c? C'est que nous *définissons* la somme de a et b justement comme c, or les définitions ne se discutent pas. Parfois ces définitions paraîtront un peu étranges, mais il n'y a rien d'étonnant, car tout ce qui est nouveau est toujours surprenant. Tout est question d'habitude, et les appareils aussi remarquables qu'un récepteur de télévision ou un téléphone n'étonnent plus personne aujourd'hui. Mais si à un élève du cours moyen habitué à ce que la somme de deux nombres a et b est le nombre d'objets contenus dans la réunion d'une collection de a objets et d'une collection de b objets (voir fig. 1 à la page 12), et le produit ab est le nombre d'objets contenus dans la réunion de a collections de b objets chacune (voir fig. 2 à la page 13), on explique ce qu'est une fraction et qu'on lui définisse la somme et le produit

des fractions $\frac{a}{c}$ et $\frac{b}{d}$ comme suit:

$$\frac{a}{c}+\frac{b}{d}=\frac{ad+bc}{cd}, \quad \frac{a}{c}\cdot\frac{b}{d}=\frac{ab}{cd},$$

cet élève sera certainement surpris par ces règles (qui nous paraissent, à nous, des plus naturelles!)

Passons maintenant aux exemples promis.

● EXEMPLE 1. *Algèbre de deux nombres.* Admettons que l'algèbre que l'on se propose d'étudier ne possède que deux éléments que, pour plus de commodité, nous appellerons nombres et désignerons par les symboles bien connus 0 et 1 (ces symboles ont ici un tout autre sens). Nous allons définir la *multiplication* de ces nombres comme une opération absolument identique à celle connue en arithmétique, c'est-à-dire à l'aide de la « table de multiplication » suivante:

·	0	1
0	0	0
1	0	1

et l'*addition* comme « d'habitude » à cette seule différence que la somme $1 + 1$ ne sera plus égale à 2 (vu que dans notre « algèbre de deux nombres » ce nombre n'existe pas!), mais de nouveau à 1. Ainsi, la « table d'addition » s'écrit

+	0	1
0	0	1
1	1	1

Nous constatons qu'en algèbre ainsi définie les deux lois commutatives sont vérifiées:

$a + b = b + a$ et $ab = ba$ quels que soient a et b.

Il est aisé de s'assurer que les lois associatives sont également satisfaites :

$$(a + b) + c = a + (b + c) \text{ et } (ab)\, c = a\, (bc) \text{ quels que soient } a,\ b \text{ et } c.$$

Soulignons que la loi associative pour la multiplication ne nécessite aucune vérification, car l'opération de multiplication que nous avons définie coïncide entièrement avec la multiplication des nombres satisfaisant à la loi associative. Il est aisé de voir que les lois idempotentes sont à leur tour vérifiées :

$$a + a = a \text{ et } aa = a \text{ quel que soit } a,$$

c'est-à-dire pour $a = 0$ et $a = 1$ (ce qui explique pourquoi nous avons posé $1 + 1 = 1$!). Les lois distributives

$$(a + b)\, c = ac + bc \text{ et } ab + c = (a + c)\,(b + c) \text{ quels que soient } a,\ b,\ c,$$

sont plus difficiles à vérifier. Ainsi, par exemple, on a

$$(1 + 1) \cdot 1 = 1 \cdot 1 = 1 \text{ et } (1 \cdot 1) + (1 \cdot 1) = 1 + 1 = 1;$$
$$(1 \cdot 1) + 1 = 1 + 1 = 1 \text{ et } (1 + 1) \cdot (1 + 1) = 1 \cdot 1 = 1.$$

Si, enfin, on convient d'attribuer le rôle de l'élément O de l'algèbre de deux nombres au nombre 0, et celui de l'élément I au nombre 1, alors les règles relatives aux éléments « particuliers » O et I seront vérifiées toujours (c'est-à-dire aussi bien pour $a = 0$, que pour $a = 1$)

$$a + 0 = a \text{ et } a \cdot 1 = a;\ a + 1 = 1 \text{ et } a \cdot 0 = 0.$$

● EXEMPLE 2. *Algèbre de quatre nombres.* Voici encore un exemple analogue, mais un peu plus compliqué. Admettons que notre algèbre comprenne quatre « nombres » que nous désignerons par les chiffres 0 et 1 et par les lettres p et q.

Définissons l'*addition* et la *multiplication* dans l'algèbre considérée à l'aide de tableaux suivants:

$+$	0	p	q	1
0	0	p	q	1
p	p	p	1	1
q	q	1	q	1
1	1	1	1	1

$\cdot$	0	p	q	1
0	0	0	0	0
p	0	p	0	p
q	0	0	q	q
1	0	p	q	1

Dans ce cas aussi on se convainc aisément que:

$a + b = b + a$ et $ab = ba$ quels que soient a, b;

$(a + b) + c = a + (b + c)$ et $(ab)\,c = a\,(bc)$ quels que soient a, b, c;

$(a + b)\,c = ac + bc$ et $ab + c = (a + c)(b + c)$ quels que soient a, b, c;

$a + a = a$ et $aa = a$ quel que soit a (c'est-à-dire pour $a = 0, p, q, 1$).

D'autre part, les nombres 0 et 1 jouent ici le rôle des éléments O et I de l'algèbre des ensembles, car pour tout a on a

$$a + 0 = a \text{ et } a \cdot 1 = a; \quad a + 1 = 1 \text{ et } a \cdot 0 = 0.$$

● EXEMPLE 3. *Algèbre des maxima et des minima.* Prenons pour éléments de notre algèbre un ensemble quelconque (borné) des nombres. Convenons de considérer comme éléments de l'algèbre des nombres quelconques x (peut-être *tous*!), où $0 \leqslant x \leqslant 1$, c'est-à-dire les nombres de l'intervalle (0, 1) bornes comprises. Nous définirons l'addition et la multiplication d'une tout autre façon: afin de ne pas confondre ces opérations avec l'addition et la multiplication ordinaires, nous les désignerons par les nouveaux symboles: $\oplus$ pour l'addition et $\otimes$ pour la multiplication. Nous conviendrons de considérer que la *somme* $x \oplus y$ de deux nombres

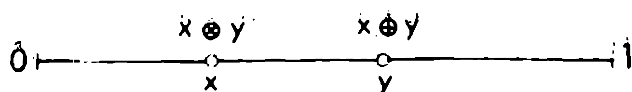


Fig. 11

x et y est égale *au plus grand* d'entre eux (ou à l'un des deux si $x = y$); par *produit* $x \otimes y$ de deux nombres x et y on entendra *le plus petit* d'entre eux (ou l'un des deux si $x = y$). Si les éléments de notre algèbre sont par exemple les nombres $0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}$ et 1, le « tableau d'addition » et le « tableau de multiplication » s'écrivent alors:

$\oplus$	0	1/3	1/2	2/3	1
0	0	1/3	1/2	2/3	1
1/3	1/3	1/3	1/2	2/3	1
1/2	1/2	1/2	1/2	2/3	1
2/3	2/3	2/3	2/3	2/3	1
1	1	1	1	1	1

$\otimes$	0	1/3	1/2	2/3	1
0	0	0	0	0	0
1/3	0	1/3	1/3	1/3	1/3
1/2	0	1/3	1/2	1/2	1/2
2/3	0	1/3	1/2	2/3	2/3
1	0	1/3	1/2	2/3	1

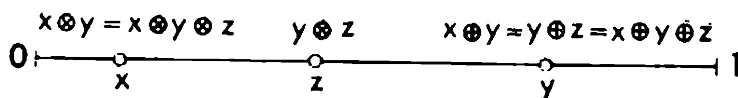
Dans les mathématiques le plus grand de deux ou de plusieurs nombres $u, v, \ldots, z$ est souvent désigné par le symbole max $[u, v, \ldots, z]$ et le plus petit de ces nombres par le symbole min $[u, v, \ldots, z]$. Ainsi, en « algèbre des maxima et des minima » on a par définition

$$x \oplus y = \max [x, y] \quad \text{et} \quad x \otimes y = \min [x, y].$$

On peut de même convenir de désigner les nombres par les points de l'axe numérique; alors les nombres x, où $0 \leqslant x \leqslant 1$, seront représentés par les points d'un segment horizontal de longueur 1, et la somme $x \oplus y$ de deux nombres x et y sera représentée par celui des points x, y qui est *à droite* et leur produit $x \otimes y$ par le point qui est *à gauche* (voir fig. 11).

Il est clair que les nouvelles opérations d'addition et de multiplication satisfont aux lois commutatives:

$$x \oplus y = y \oplus x \quad \text{et} \quad x \otimes y = y \otimes x,$$

Fig. 12 

de même qu'aux lois associatives:

$$(x \oplus y) \oplus z = x \oplus (y \oplus z) \text{ et } (x \otimes y) \otimes z = x \otimes (y \otimes z);$$

ainsi, par exemple, le nombre $(x \oplus y) \oplus z$ ou $x \oplus (y \oplus z)$ (ce dernier peut tout simplement être désigné par $x \oplus y \oplus z$, sans parenthèses) représente le max $[x, y, z]$ (voir fig. 12) et le nombre $(x \otimes y) \otimes z$ ou $x \otimes (y \otimes z)$ (ou tout simplement $x \otimes y \otimes z$, sans parenthèses) le min $[x, y, z]$ (voir fig. 12). Il n'est pas moins clair que les lois idempotentes

$$x \oplus x = \max [x, x] = x \quad \text{et} \quad x \otimes x = \min [x, x] = x$$

ont également lieu.

Vérifions, enfin, la validité des lois distributives:

$$(x \oplus y) \otimes z = (x \otimes z) \oplus (y \otimes z)$$

et

$$(x \otimes y) \oplus z = (x \oplus z) \otimes (y \oplus z).$$

Il est clair que le nombre

$$(x \oplus y) \otimes z = \min\{\max [x, y], z\}$$

est égal à z si *l'un au moins des nombres* x, y *est supérieur à* z et égal au plus grand de ces nombres si *ces nombres sont inférieurs à* z (voir fig. 13,*a*, *b*). Il en est de même pour le nombre

$$(x \otimes z) \oplus (y \otimes z) = \max\{\min [x, z], \min [y, z]\}$$

(voir fig. 13). D'une façon analogue, le nombre

$$(x \otimes y) \oplus z = \max\{\min [x, y], z\}$$

est égal à z si *l'un au moins des nombres* x, y *est inférieur à* z et égal au plus petit des nombres x, y s'*ils sont tous les deux*

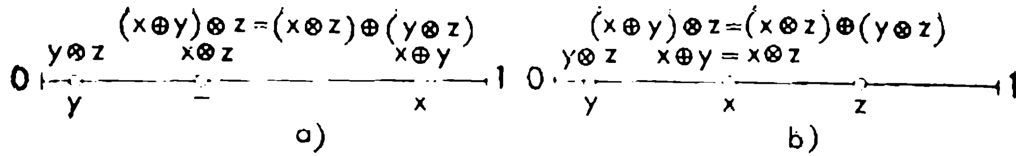


Fig. 13

supérieurs à z (voir fig. 14,*a*, *b*). Il en est de même pour le nombre

$$(x \oplus z) \otimes (y \oplus z) = \min\{\max[x, z], \max[y, z]\}$$

(voir fig. 14).

Pour voir si *toutes* les lois de l'algèbre des ensembles sont vérifiées dans notre algèbre originale, il suffit de noter que les rôles des éléments O et I de l'algèbre des ensembles y sont assumés par le plus petit des nombres considérés, c'est-à-dire le nombre 0, et par le nombre le plus grand des nombres considérés, c'est-à-dire le nombre 1. En effet, quel que soit x, $0 \leqslant x \leqslant 1$, on a toujours

$$x \oplus 0 = \max[x, 0] = x \quad \text{et} \quad x \otimes 1 = \min[x, 1] = x;$$
$$x \oplus 1 = \max[x, 1] = 1 \quad \text{et} \quad x \otimes 0 = \min[x, 0] = 0.$$

● EXEMPLE 4. *Algèbre des plus petits multiples et des plus grands diviseurs.* Soit N un nombre entier positif quelconque. Prenons pour éléments de notre nouvelle algèbre tous les *diviseurs* possibles du nombre N. Si par exemple $N = 210 = 2 \cdot 3 \cdot 5 \cdot 7$, les éléments de l'algèbre considérée seront 1, 2, 3, 5, 6, 7, 10, 14, 15, 21, 30, 35, 42, 70, 105 et 210. Voici comment nous allons définir l'addition et la multiplication de ces nombres: par la *somme* $m \oplus n$ des nombres m et n nous entendrons *leur plus petit commun multiple* (P.P.C.M.), c'est-à-dire le plus petit nombre entier (positif) qui est divisible par les deux nombres m et n; par le *produit* $m \otimes n$ des nombres m et n nous entendrons *le plus grand commun diviseur* (P.G.C.D.) de ces nombres, c'est-à-dire le plus grand nombre entier qui divise

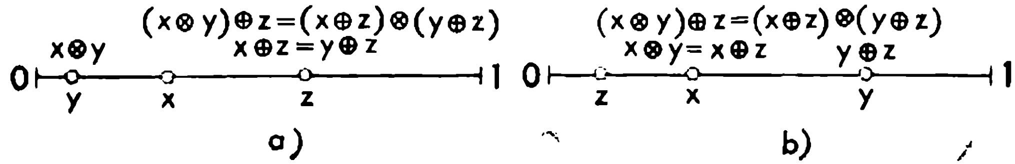


Fig. 14

m et *n*. Ainsi, par exemple, si $N = 6$ et notre algèbre ne contient que quatre nombres 1, 2, 3 et 6, l'addition et la multiplication s'écrivent :

$\oplus$	1	2	3	6
1	1	2	3	6
2	2	2	6	6
3	3	6	3	6
6	6	6	6	6

et

$\otimes$	1	2	3	6
1	1	1	1	1
2	1	2	1	2
3	1	1	3	3
6	1	2	3	6

En théorie des nombres on désigne souvent le P.P.C.M. de deux ou de plusieurs nombres $m, n, \ldots, s$ par $[m, n, \ldots, s]$ et le P.G.C.D. par $(m, n, \ldots, s)$. De cette façon, dans notre algèbre on a par définition

$$m \oplus n = [m, n] \quad \text{et} \quad m \otimes n = (m, n).$$

Si l'algèbre contient, par exemple, les nombres 10 et 15, on a

$$10 \oplus 15 = [10, 15] = 30 \quad \text{et} \quad 10 \otimes 15 = (10, 15) = 5.$$

Il est évident que dans notre algèbre on aura toujours :

$$m \oplus n = n \oplus m \quad \text{et} \quad m \otimes n = n \otimes m.$$

En poursuivant, on peut donc écrire :

$$(m \oplus n) \oplus p = m \oplus (n \oplus p) \qquad (= [m, n, p])$$

(on peut d'ailleurs convenir d'écrire ce nombre $m \oplus n \oplus p$, sans parenthèses), et

$$(m \otimes n) \otimes p = m \otimes (n \otimes p) \quad (= (m, n, p))$$

(ou $m \otimes n \otimes p$). Les lois idempotentes sont aussi évidentes :

$$m \oplus m = [m, m] = m \quad \text{et} \quad m \otimes m = (m, m) = m.$$

La vérification des lois distributives est (comme toujours d'ailleurs) plus difficile à réaliser. Le nombre

$$(m \oplus n) \otimes p = ([m, n], p)$$

n'est rien d'autre que le *P.G.C.D. du nombre p et du P.P.C.M. de m et n* (pensez-y bien !) ; ce nombre ne contient que ceux des facteurs premiers qui font partie de p et, en même temps, de l'*un au moins* des nombres m et n. Il est évident que ces facteurs premiers (et seulement eux) entrent également dans la décomposition du nombre

$$(m \otimes p) \oplus (n \otimes p) = [(m, p), (n, p)] ;$$

on a donc toujours

$$(m \oplus n) \otimes p = (m \otimes p) \oplus (n \otimes p).$$

Si, par exemple, on puise la réserve des nombres à partir de l'ensemble des diviseurs du nombre 210, il vient

$$(10 \oplus 14) \otimes 105 = ([10, 14], 105) = (70, 105) = 35$$

et

$$(10 \otimes 105) \oplus (14 \otimes 105) = [(10, 105), (14, 105)] = [5, 7] = 35.$$

D'une manière analogue, le nombre

$$(m \otimes n) \oplus p = [(m, n), p]$$

qui est le *P.P.C.M. du nombre p et le P.G.C.D. de m et n*, ne contient que ceux des facteurs premiers qui entrent ou bien dans la décomposition de p, ou bien dans la décomposition des nombres m et n (ou, peut être, simultanément de p et des deux nombres m et n). Mais le nombre

$$(m \oplus p) \otimes (n \oplus p) = ([m, p], [n, p])$$

contient exactement ces mêmes facteurs; pour cette raison on a toujours

$$(m \otimes n) \oplus p = (m \oplus p) \otimes (n \oplus p).$$

Ainsi,

$$(10 \otimes 14) \oplus 105 = [(10, 14), 105] = (2, 105) = 210$$

et

$$(10 \oplus 105) \otimes (14 \oplus 105) = ([10, 105], [14, 105]) = \\ = (210, 210) = 210.$$

Enfin, le rôle des éléments O et I de l'algèbre des ensembles revient ici au nombre 1, qui est le plus petit des nombres considérés, et au nombre N, qui en est le plus grand. En effet, il est évident (n'oublions pas que notre algèbre n'est formée que des diviseurs du nombre N) que

$$m \oplus 1 = [m, 1] = m \quad \text{et} \quad m \otimes N = (m, N) = m;$$
$$m \oplus N = [m, N] = N \quad \text{et} \quad m \otimes 1 = (m, 1) = 1.$$

De cette façon, ici également les lois de l'algèbre des ensembles sont toutes vérifiées.

Ainsi, nous voyons qu'il existe une quantité de systèmes divers d'« objets » (qui sont les *éléments* de l'algèbre étudiée) pour lesquels on peut définir les opérations d'*addition* et de *multiplication*, qui obéissent à toutes les règles connues de l'algèbre des ensembles: deux lois commutatives, deux lois associatives, deux lois distributives, deux lois idempotentes et, enfin, quatre règles qui définissent les propriétés d'éléments « particuliers » assumant dans notre algèbre le rôle proche de celui du zéro et de l'unité. Dans ce qui suit nous examinerons encore deux exemples particulièrement importants et intéressants de telles algèbres.

Avant de passer à l'étude détaillée des propriétés générales de toutes ces algèbres, il nous faut avant tout les réunir sous un nom générique quelconque. Etant donné que les algèbres aux propriétés si étranges furent examinées pour la première fois par le remarquable mathématicien anglais

du XIX siècle George Boole, on les appelle aujourd'hui ***algèbres de Boole*** *). Pour les opérations principales de l'algèbre de Boole, nous conserverons toujours les termes adoptés d'« addition » et de « multiplication » (mais sans oublier toutefois que ces opérations n'ont rien de commun avec les opérations sur les nombres!) ; cependant nous les appellerons parfois « *addition* et *multiplication*» *booléennes.*

L'ouvrage de G. Boole consacré à l'étude détaillée de l'algèbre non ordinaire qui fait objet du présent article a vu son appartion en 1854, c'est-à-dire il y a plus de 100 ans sous le titre : « An Investigation of the laws of thought ». Pour le moment ce titre peut paraître étrange, mais après avoir lu ce livre le lecteur comprendra le lien qui existe entre les algèbres non ordinaires passées ici en revue et les lois de la pensée. Il est à remarquer que ce sont précisément les liens des algèbres de Boole avec les « lois de la pensée » qui ont éveillé un intérêt tout particulier de la part des mathématiciens pour ces œuvres passées d'abord inaperçues. Ces dernières années l'ouvrage de Boole a été maintes fois publié et traduit en plusieurs langues et la notion même de l'algèbre de Boole est devenue partie intégrante du programme scolaire. L'enseignement de l'algèbre de Boole dans le secondaire est à l'étude dans nombre de pays. D'ores et déjà il fait l'objet de débats entre les mathématiciens et les pédagogues.

Exercices

1. Vérifier directement que les deux lois distributives sont vraies pour les trois éléments de l'« algèbre de Boole de deux nombres » (exemple 1, p. 31).

2. Vérifier que les deux lois distributives sont vraies pour plusieurs combinaisons de trois éléments de l'« algèbre de Boole de quatre nombres » (exemple 2, p. 32).

---

*) Pour la définition exacte des algèbres de Boole, voir l'Annexe (p. 90).

3. a) Si, par exemple, dans un appartement quelconque il n'y a qu'un seul écolier, alors les « ensembles des écoliers se trouvant dans cet appartement » seront les suivants: l'ensemble I composé d'un seul écolier et l'ensemble O ne possédant aucun écolier (l'ensemble vide). On demande de dresser pour l'« algèbre des ensembles des écoliers qui habitent cet appartement » (celle-ci ne possède que deux éléments: O et I) le « tableau d'addition » et le « tableau de multiplication » et de les comparer ensuite avec ceux mentionnés à la page 31 ; en déduire que l'« algèbre de deux nombres » (exemple 1 de ce paragraphe) obéit à toutes les lois de l'algèbre de Boole.

b) Soient deux écoliers (Pierre et Catherine) habitant un même appartement. Alors l'« algèbre des ensembles des écoliers habitant cet appartement » comprendra quatre éléments: l'ensemble I des deux écoliers; les deux ensembles P (Pierre) et C (Catherine) composés chacun d'un seul écolier; l'ensemble vide O. Dresser le « tableau d'addition » et le « tableau de multiplication » pour cette algèbre des ensembles et les comparer avec ceux de la page 32 ; en déduire que toutes les lois de l'algèbre de Boole sont applicables à l'« algèbre de quatre nombres » de l'exemple 2.

4. Vérifier que

a) $\min\{\max[1/2,\ 1/3],\ 1/4\} = \max\{\min[1/2,\ 1/4],\ \min[1/3,\ 1/4]\}$

et $\max\{\min[1/2,\ 1/3],\ 1/4\} = \min\{\max[1/2,\ 1/4],\ \max[1/3,\ 1/4]\}$.

b) $([12,\ 30],\ 8) = [(12,\ 8),\ (30,\ 8)]$ et $[(12,\ 30),\ 8] = ([12,\ 8],\ [30,\ 8])$.

5. a) Former le « tableau d'addition » et le « tableau de multiplication » de l'algèbre de Boole comprenant trois nombres 0, 1/2 et 1, où $x \oplus y = \max[x, y]$ et $x \otimes y = \min[x, y]$; s'assurer que les lois de l'algèbre de Boole sont satisfaites.

b) Former le « tableau d'addition » et le « tableau de multiplication » de l'algèbre des diviseurs du nombre 12, où $m \oplus n = [m, n]$ et $m \otimes n = (m, n)$; vérifier que certaines lois de l'algèbre de Boole sont applicables à l'algèbre ainsi définie.

6*. Soit la décomposition d'un nombre N (entier positif) en facteurs premiers

$$N = p_1^{A_1} p_2^{A_2} \dots p_k^{A_k} ;$$

alors deux diviseurs m et n quelconques de ce nombre peuvent être mis sous la forme

$$m = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k},$$

où $0 \leqslant a_1 \leqslant A_1,\quad 0 \leqslant a_2 \leqslant A_2, \dots,\quad 0 \leqslant a_k \leqslant A_k,$

et

$n = p_1^{b_1} p_2^{b_2} \dots p_k^{b_k}$,

où $0 \leqslant b_1 \leqslant A_1, \quad 0 \leqslant b_2 \leqslant A_2, \dots, \quad 0 \leqslant b_k \leqslant A_k$,

(parmi les nombres $a_1, a_2, \dots, a_k$; $b_1, b_2, \dots, b_k$ il y a ceux qui peuvent être nuls). Décomposer les nombres $[m, n]$ (le P.P.C.M. des nombres m et n) et (m, n) (le P.G.C.D. de m et n) en facteurs premiers. Démontrer, en utilisant la décomposition ci-dessus, que l'ensemble de tous les diviseurs du nombre N muni des opérations $m \oplus n = [m, n]$ et $m \otimes n = (m, n)$ forme une algèbre de Boole.

§ 3. AUTRES PROPRIÉTÉS DES ALGÈBRES DE BOOLE : PRINCIPE DE DUALITÉ ; ÉGALITÉS ET INÉGALITÉS BOOLÉENNES

Poursuivons l'étude de l'algèbre de Boole. Ce qui saute avant tout aux yeux, c'est le parallélisme entre les propriétés des opérations d'addition et de multiplication de cette algèbre : ces opérations sont à tel point identiques que dans *toute formule* (correcte, bien sûr !) *de l'algèbre de Boole on peut remplacer l'addition par la multiplication et inversement.* Ainsi, comme il a été démontré plus haut, l'égalité suivante de l'algèbre de Boole

$$A (A + C) (B + C) = AB + AC$$

est vérifiée (voir l'exemple examiné à la fin des exercices du § 1, p. 27). En remplaçant dans cet exemple l'addition par la multiplication et vice versa, on obtient l'égalité

$$A + AC + BC = (A + B) (A + C),$$

qui est également juste (cf. page 44). Soulignons que si les *éléments « particuliers » O et I figurent dans une égalité quelconque de l'algèbre de Boole, alors en y remplaçant l'addition par la multiplication et vice versa, il faut en même temps remplacer O par I et inversement.* Ainsi, l'égalité

$$(A + B) (A + I) + (A + B) (B + O) = A + B$$

(cf. exercice 8 de la page 28) entraîne l'égalité suivante

$$(AB + AO)(AB + BI) = AB.$$

La propriété énoncée de l'algèbre de Boole qui permet d'obtenir (sans démonstration à partir d'une égalité quelconque) une égalité nouvelle *) s'appelle *principe de dualité*, et les égalités elles-mêmes qui s'obtiennent l'une de l'autre selon ce principe *égalités duales*. Le principe de dualité découle du fait que les lois fondamentales de l'algèbre de Boole qui nous servent à démontrer telle ou telle *formule de Boole* sont absolument symétriques: chaque loi de cette algèbre possède une loi duale, c'est-à-dire une loi que l'on obtient de la loi initiale en remplaçant l'addition par la multiplication et inversement, l'élément O par I et inversement. Ainsi, la loi commutative de l'addition a pour duale la loi de la multiplication, la loi associative de l'addition

---

*) Cette « nouvelle » égalité que l'on obtient à partir d'une formule quelconque de l'algèbre de Boole en y remplaçant l'addition par la multiplication et inversement peut parfois se confondre avec l'égalité initiale, dans ce cas-là ce procédé ne nous donne aucun avantage. Si, par exemple, dans l'égalité (ex. 6, page 28)

$$(A + B)(B + C)(C + A) = AB + BC + CA$$

on remplace l'addition par la multiplication et inversement on obtient l'égalité

$$AB + BC + CA = (A + B)(B + C)(C + A),$$

qui coïncide avec l'égalité initiale, quant à l'égalité (ex. 7, page 28)

$$(A + B)(B + C)(C + D) = AC + BC + BD$$

si l'on y remplace la multiplication par l'addition et inversement on obtient l'égalité

$$AB + BC + CD = (A + C)(B + C)(B + D)$$

qui diffère de l'égalité initiale de façon insignifiante (ces deux égalités deviennent absolument identiques si l'on remplace dans cette dernière B par C et C par B).

a pour duale la loi de la multiplication, la loi idempotente de l'addition a pour duale la loi de la multiplication; la première loi distributive a pour duale la deuxième loi distributive et, enfin, aux égalités $A + O = A$ et $A + I = I$ correspondent respectivement les égalités $AI = A$ et $AO = O$. De cette façon, si, pour démontrer une égalité quelconque, l'on se sert des lois fondamentales de l'algèbre de Boole, alors les lois duales nous permettront de démontrer l'égalité qui est duale de la première.

● EXEMPLE. Soit à démontrer l'égalité

$$A + AC + BC = (A + B)(A + C)$$

duale de

$$A(A + C)(B + C) = AB + AC.$$

En effet,

$$A + AC + BC \underset{\substack{\text{associativité}\\ \text{de l'addition}}}{=} A + (AC + BC) \underset{\substack{\text{1ère loi}\\ \text{distributive}}}{=} A + (A + B)C =$$

$$\underset{\substack{\text{commutativité}\\ \text{de l'addition}}}{=} (A + B)C + A \underset{\substack{\text{2ème loi}\\ \text{distributive}}}{=} [(A + B) + A](C + A) =$$

$$\underset{\substack{\text{commutativité et}\\ \text{associativité}\\ \text{de l'addition}}}{=} [(A + A) + B](A + C) \underset{\substack{\text{idempotence}\\ \text{de l'addition}}}{=} (A + B)(A + C)$$

(cf. avec la démonstration de l'égalité $A(A + C)(B + C) = AB + AC$ à la page 28).

Pour la démonstration du principe de dualité, l'algèbre de Boole possède une opération toute particulière qui permet de transformer chaque élément A de cette algèbre en un élément nouveau $\bar{A}$ *en remplaçant l'addition par la multiplication et inversement*. En d'autres termes, cette opération (que nous appellerons *opération « barre »*) est telle que

$$\overline{A + B} = \bar{A}\bar{B} \quad \text{et} \quad \overline{AB} = \bar{A} + \bar{B}.$$

D'autre part

$$\bar{O} = I \quad \text{et} \quad \bar{I} = O.$$

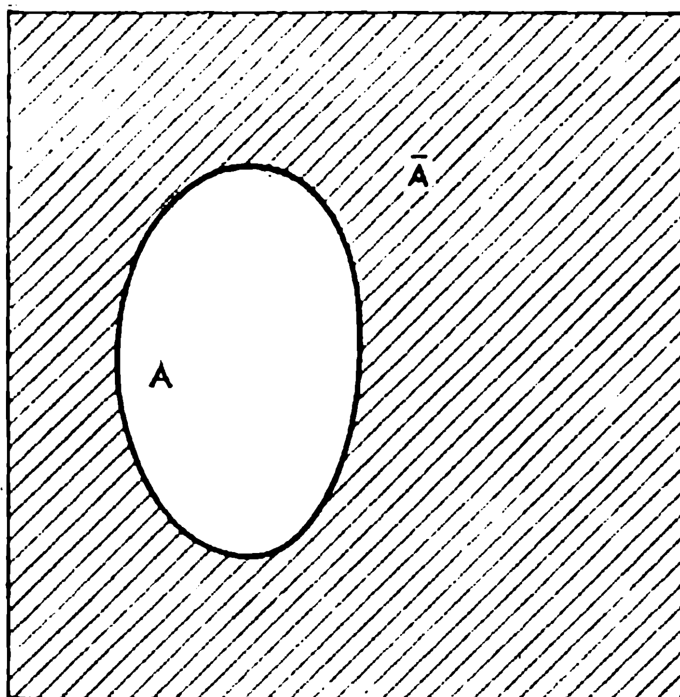


Fig. 15

L'opération « barre » transforme l'élément $\bar{A}$ en l'élément initial A, c'est-à-dire pour chaque élément A de l'algèbre de Boole on a

$$\bar{\bar{A}} = \overline{(\bar{A})} = A.$$

Dans l'*algèbre des ensembles*, l'opération « barre » (cette opération originale permet de former un élément nouveau de l'algèbre de Boole non pas à partir de deux éléments connus comme dans le cas de l'addition et de la multiplication, mais à partir d'un *seul*!) peut être interprétée comme suit. Par ensemble $\bar{A}$ on entend le *complément* de l'ensemble A, c'est-à-dire un *ensemble composé uniquement de tels éléments de l'ensemble universel I qui n'appartiennent pas à l'ensemble A* (fig. 15). Ainsi, par exemple, si l'on prend pour l'ensemble universel l'ensemble de tous les élèves d'une classe et pour A l'ensemble des élèves qui ont obtenu au moins une mauvaise note (l'ensemble des élèves faibles), alors $\bar{A}$ est l'ensemble des élèves qui ont obtenu au moins

« assez bien » dans toutes les disciplines (l'ensemble des bons élèves).

De la définition même du complément $\bar{A}$ de l'ensemble A il s'ensuit que

$$\bar{\bar{A}} = \overline{(\bar{A})} = A$$

et

$$A + \bar{A} = I, \quad A\bar{A} = O$$

(voir fig. 15 ; les deux dernières égalités peuvent même servir de *définition* de l'ensemble $\bar{A}$). Il est évident que

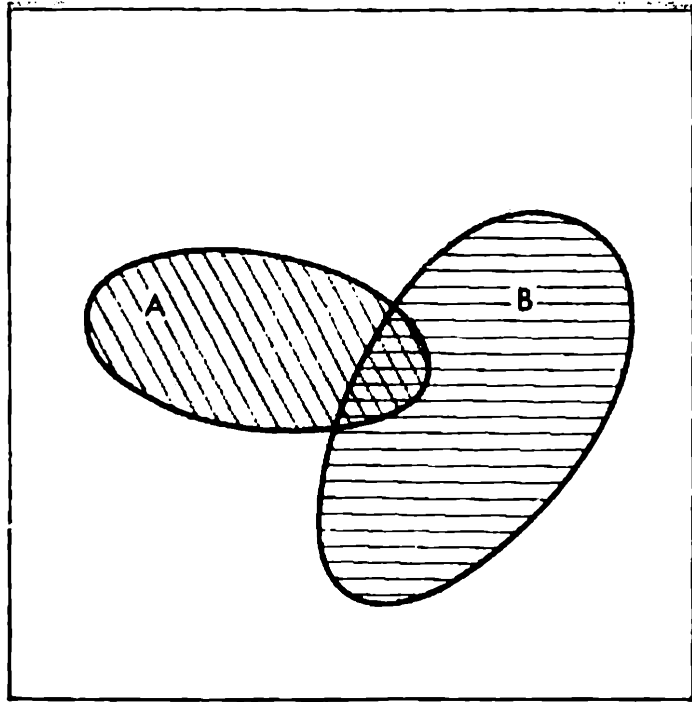
$$\bar{O} = I, \quad \bar{I} = O.$$

Démontrons, enfin, que les propriétés les plus importantes de l'opération « barre » sont vérifiées dans l'algèbre des ensembles :

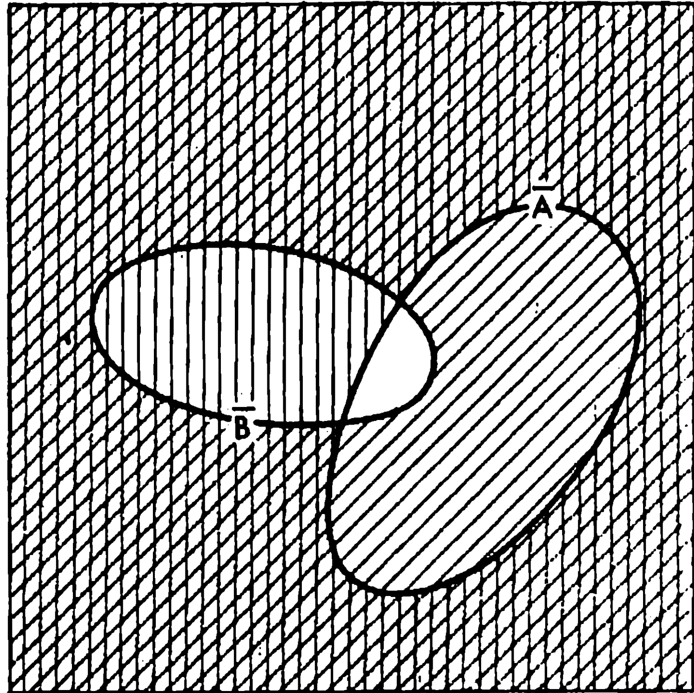
$$\overline{A + B} = \bar{A}\bar{B} \quad \text{et} \quad \overline{AB} = \bar{A} + \bar{B};$$

ces règles sont appelées *règles de De Morgan*, du nom du mathématicien anglais Augustus de Morgan (1806-1871), contemporain et confrère de George Boole. La figure 16,*a* représente l'ensemble A marqué de hachures inclinées à gauche et la figure 16,*b* son complément $\bar{A}$ relativement au carré I marqué de hachures inclinées à droite ; la figure 16,*a* représente l'ensemble B marqué de hachures horizontales et la figure 16,*b* son complément $\bar{B}$ marqué de hachures verticales. De cette façon c'est l'ensemble $\overline{A + B}$ qui se trouve haché sur la figure 16,*a*, et c'est l'ensemble $\bar{A}\bar{B}$ qui est doublement haché sur la figure 16,*b*. Il résulte de la comparaison des figures 16,*a* et 16,*b* que l'ensemble $\bar{A}\bar{B}$ complète l'ensemble $\overline{A + B}$ de la figure 16,*a*, ce qui démontre la première règle de De Morgan :

$$\overline{A + B} = \bar{A}\bar{B}.$$



a)



b)

Fig. 16

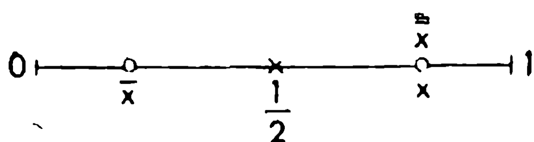


Fig. 17

D'autre part, l'ensemble AB représenté sur la figure 16,*a* par des hachures entrecroisées est le complément de l'ensemble $\bar{A} + \bar{B}$ de la figure 16,*b*. On a donc

$$\overline{AB} = \bar{A} + \bar{B}.$$

Elucidons maintenant le sens de l'opération « barre » pour les autres exemples d'algèbres de Boole examinés plus haut. Ainsi, dans l'*algèbre de deux nombres* (exemple 1, p. 31) on a

$$\bar{0} = 1, \quad \bar{1} = 0.$$

On comprend aisément que pour tout élément a de notre algèbre (c'est-à-dire pour $a = 0$ et pour $a = 1$) on a $\bar{\bar{a}} = a$. Ensuite il résulte de la comparaison du « tableau d'addition » et du « tableau de multiplication », dressé pour les nombres $\bar{0} = 1$ et $\bar{1} = 0$:

+	0	1
0	0	1
1	1	1

et

·	$\bar{0}=1$	$\bar{1}=0$
$\bar{0}=1$	1	0
$\bar{1}=0$	0	0

que dans tous les cas $\overline{a + b} = \bar{a}\bar{b}$. On vérifie de façon analogue la deuxième règle de De Morgan $\overline{ab} = \bar{a} + \bar{b}$.

Dans l'*algèbre de quatre nombres* (exemple 2, p. 32) on a

$$\bar{0} = 1, \quad \bar{p} = q, \quad \bar{q} = p, \quad \bar{1} = 0.$$

Il est évident que dans ce cas aussi la relation $\bar{\bar{a}} = a$ est satisfaite pour tout élément a de cette algèbre. Pour vérifier la relation $\overline{a + b} =$

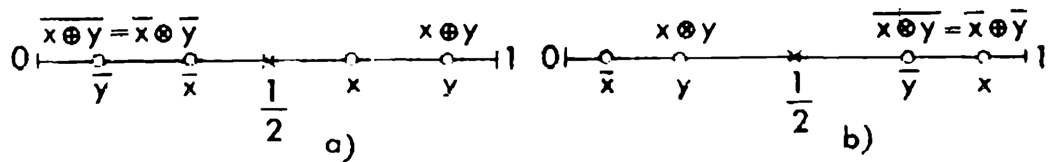


Fig. 18

$= \bar{a}\bar{b}$ il suffit de comparer les deux tableaux suivants:

$+$	0	p	q	1
0	0	p	q	1
p	p	p	1	1
q	q	1	q	1
1	1	1	1	1

$\cdot$	$\bar{0}=1$	$\bar{p}=q$	$\bar{q}=p$	$\bar{1}=0$
$\bar{0}=1$	1	q	p	0
$\bar{p}=q$	q	q	0	0
$\bar{q}=p$	p	0	p	0
$\bar{1}=0$	0	0	0	0

On vérifie analogiquement la relation $\overline{ab} = \bar{a} + \bar{b}$.

Passons maintenant à l'*algèbre des maxima et des minima*, dont les éléments sont les nombres x tels que $0 \leqslant x \leqslant 1$ et l'addition $\oplus$ et la multiplication $\otimes$ booléennes se définissent comme suit:

$$x \oplus y = \max [x, y], \quad x \otimes y = \min [x, y].$$

Pour que les règles de De Morgan soient vérifiées dans cette algèbre

$$\overline{x \oplus y} = \bar{x} \otimes \bar{y}, \quad \overline{x \otimes y} = \bar{x} \oplus \bar{y}$$

c'est-à-dire pour que l'on ait

$$\overline{\max [x, y]} = \min [\bar{x}. \bar{y}], \quad \overline{\min [x, y]} = \max [\bar{x}, \bar{y}],$$

il faut que l'opération « barre » *inverse l'ordre* des éléments, c'est-à-dire que la condition $x \leqslant y$ entraîne $\bar{x} \geqslant \bar{y}$. Si donc, on prend pour les éléments de l'algèbre *tous* les nombres x tels que $0 \leqslant x \leqslant 1$, on peut alors poser, par exemple,

$$\bar{x} = 1 - x:$$

en d'autres termes, on peut considérer que le *point* $\bar{x}$ *est symétrique du point* x *par rapport au milieu* 1/2 *du segment* [0, 1] (voir fig. 17). Il est

alors clair que

$$\bar{0}=1, \quad \bar{1}=0$$

et

$$\bar{\bar{x}}=x.$$

Il va de soi que les règles de De Morgan sont également satisfaites:

$$\overline{x \oplus y}=\bar{x} \otimes \bar{y}, \quad \overline{x \otimes y}=\bar{x} \oplus \bar{y}$$

(voir fig. 18,a, b).

Examinons, enfin, l'*algèbre des plus petits multiples et des plus grands diviseurs* dont les éléments sont tous les diviseurs possibles d'un nombre entier positif N, l'addition $\oplus$ et la multiplication $\otimes$ booléennes étant définies comme suit:

$$m \oplus n = [m, n], \quad m \otimes n = (m, n),$$

où $[m, n]$ est le plus petit commun multiple des nombres m et n, et (m, n) leur plus grand commun diviseur. Posons dans ce cas

$$\bar{m}=\frac{N}{m}:$$

ainsi, dans le cas considéré plus haut où $N = 210$ on a

$$\bar{1}=210, \quad \bar{2}=105, \quad \bar{3}=70, \quad \bar{5}=42, \quad \bar{6}=35, \quad \bar{7}=30,$$

$$\overline{10}=21, \quad \overline{14}=15, \quad \overline{15}=14, \quad \overline{21}=10, \quad \overline{30}=7, \quad \overline{35}=6,$$

$$\overline{42}=5, \quad \overline{70}=3, \quad \overline{105}=2, \quad \overline{210}=1.$$

On comprend alors aisément que

$$\bar{1}=N \quad \text{et} \quad \bar{N}=1.$$

Il est de plus évident que

$$\bar{\bar{m}}=\frac{N}{N/m}=m.$$

Les règles de De Morgan y sont évidemment satisfaites:

$$\overline{m \oplus n}=\bar{m} \otimes \bar{n} \quad \text{et} \quad \overline{m \otimes n}=\bar{m} \oplus \bar{n};$$

il s'ensuit,

$$6 \oplus 21=[6, 21]=42 \quad \text{et} \quad \bar{6} \otimes \overline{21}=35 \otimes 10=$$
$$=(35, 10)=5, \text{ tandis que } \overline{42}=5$$

et

$$6 \otimes 21=(6, 21)=3 \quad \text{et} \quad \bar{6} \oplus \overline{21}=35 \oplus 10=$$
$$=[35, 10]=70, \text{ tandis que } \bar{3}=70.$$

Nous laissons au lecteur le soin de démontrer les règles de De Morgan (voir d'ailleurs l'exercice n° 6, p. 59).

Soit maintenant une égalité quelconque valable dans toute algèbre de Boole, par exemple, l'égalité suivante qui nous est déjà familière

$$A\,(A + C)\,(B + C) = AB + AC.$$

Si l'on applique maintenant aux deux membres de cette égalité l'opération « barre », on obtient

$$\overline{A\,(A+C)\,(B+C)} = \overline{AB+AC}.$$

Mais, en vertu des règles de De Morgan, on peut écrire

$$\overline{A\,(A+C)\,(B+C)} = \overline{[A\,(A+C)]\,(B+C)} =$$

$$= \overline{A\,(A+C)} + \overline{B+C} = \bar{A} + \overline{A+C} + \bar{B}\bar{C} = \bar{A} + \bar{A}\bar{C} + \bar{B}\bar{C}$$

et

$$\overline{AB+AC} = \overline{AB}\cdot\overline{AC} = (\bar{A}+\bar{B})\,(\bar{A}+\bar{C}).$$

On a finalement

$$\bar{A} + \bar{A}\bar{C} + \bar{B}\bar{C} = (\bar{A}+\bar{B})\,(\bar{A}+\bar{C}).$$

Mais comme cette égalité est vraie *quels que soient* $\bar{A}$, $\bar{B}$ et $\bar{C}$, elle le sera si l'on remplace les éléments $\bar{A}$, $\bar{B}$ et $\bar{C}$ de l'algèbre de Boole considérée par les lettres A, B et C, ce qui nous amène à l'égalité

$$A + AC + BC = (A + B)\,(A + C)$$

duale de l'égalité initiale.

Voilà comment le principe de dualité découle des propriétés de l'opération « barre » (surtout des règles de De Morgan)! N'oublions pas seulement que si l'égalité initiale comprenait les éléments « particuliers » O et I, alors en vertu de

$$\bar{O} = I \quad \text{et} \quad \bar{I} = O$$

dans l'égalité transformée (duale) au lieu de l'élément O on aurait I et inversement.

En appliquant l'opération « barre » aux deux membres de l'égalité

$$A\,(A+I)\,(B+O) = AB$$

(cf. exercice 5, p. 28) on aura

$$\overline{A\,(A+I)\,(B+O)} = \overline{AB}$$

ou, vu que

$$\overline{A\,(A+I)\,(B+O)} = \overline{A\,(A+I)} + \overline{B+O} =$$
$$= \bar{A} + \overline{A+I} + \overline{B+O} = \bar{A} + \bar{A}\bar{I} + \bar{B}\,\bar{O} = \bar{A} + \bar{A}\,\bar{O} + \bar{B}\bar{I}$$

et

$$\overline{AB} = \bar{A} + \bar{B},$$

cette égalité prend la forme

$$\bar{A} + \bar{A}O + \bar{B}I = \bar{A} + \bar{B}.$$

Mais cette dernière égalité (où $\bar{A}$ et $\bar{B}$ sont arbitraires!) est équivalente à l'égalité suivante:

$$A + AO + BI = A + B,$$

obtenue à partir de l'égalité initiale si l'on y *remplace la somme par le produit et inversement, et l'élément O par I et inversement.*

Il est à remarquer que le principe de dualité peut être appliqué non seulement aux égalités mais aussi aux « inégalités booléennes ». Mais avant de l'expliquer, voyons encore un concept qui joue un rôle fort important dans l'algèbre de Boole.

Dans chaque algèbre de Boole outre la notion d'égalité de ses éléments (l'égalité $A = B$ signifie que A et B sont tout simplement *un seul et même* élément de l'algèbre de Boole!) il existe encore une relation importante dont le rôle est identique en quelque sorte à celui joué par la relation « plus grand » (ou « plus petit ») dans l'algèbre des nombres. Cette relation est désignée par le symbole $\supset$ (ou $\subset$) et notée comme suit

$$A \supset B \text{ ou } B \subset A$$

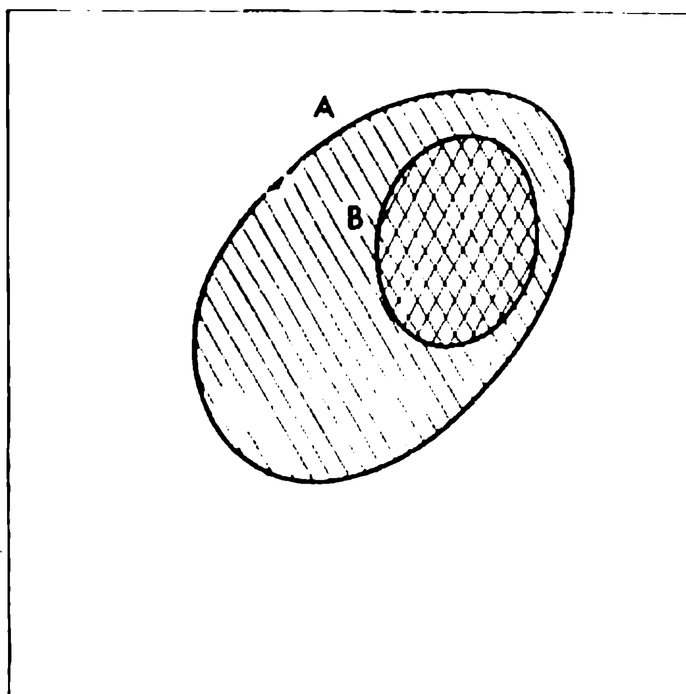


Fig. 19

(ces deux dernières corrélations ont le même sens; il est à remarquer que par leur forme elles rappellent des écritures $a > b$ ou $b < a$); dans l'algèbre des ensembles $A \supset B$ signifie que l'*ensemble A contient l'ensemble B* (voir fig. 19). Si par exemple A_2 est l'ensemble des nombres pairs et A_6 l'ensemble des nombres entiers multiples de 6, il est alors clair que $A_2 \supset A_6$; de façon analogue, si A est l'ensemble des bons élèves d'une classe et B l'ensemble des excellents élèves, il est aisé de voir que $A \supset B$. Il faut cependant tenir compte du fait que si les ensembles A et B *se confondent*, nous écrirons toujours $A \supset B$: car dans ce cas aussi l'ensemble B est entièrement contenu dans l'ensemble A ! De cette façon la relation $\supset$ pour les éléments de l'algèbre de Boole est plus proche non pas de la relation $>$ (« plus grand ») pour les nombres mais de la relation $\geqslant$ (« plus grand ou égal »).

Il est aisé de voir que

$$\text{si } A \supset B \quad \text{et} \quad B \supset C \text{ alors } A \supset C$$

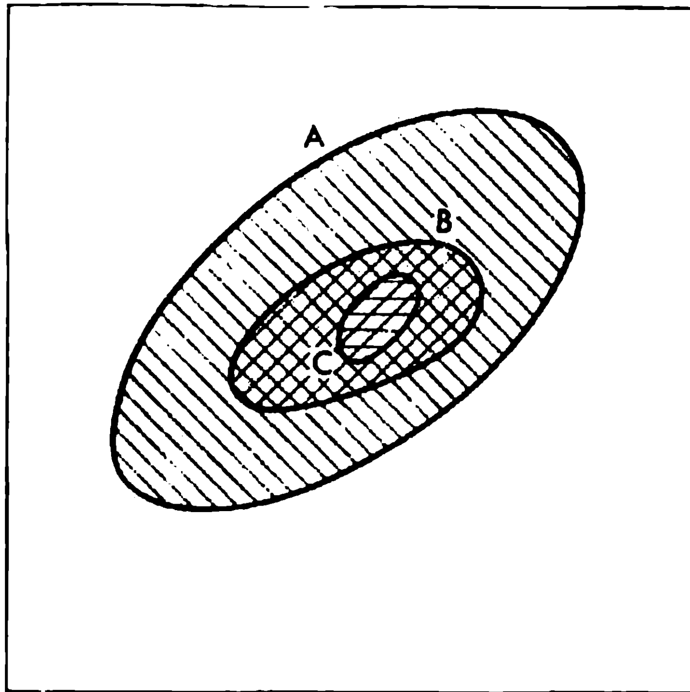


Fig. 20

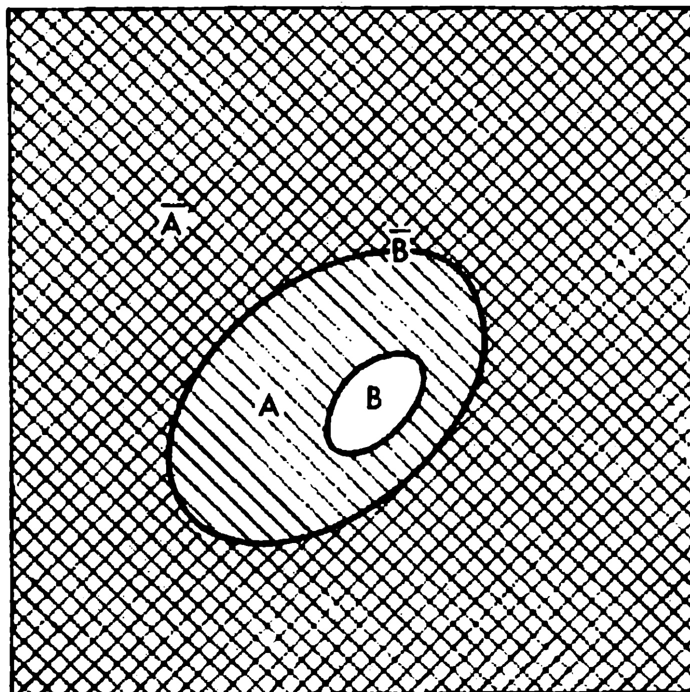


Fig. 21

(fig. 20) ; d'une façon analogue, il résulte des relations $a \geqslant b$ et $b \geqslant c$ pour les nombres, que $a \geqslant c$. Ensuite,

$$\text{si } A \supset B \text{ et } B \supset A \text{ alors } A = B,$$

de même les relations $a \geqslant b$ et $b \geqslant a$ pour les nombres entraînent $a = b$. Et, enfin, on a

$$\text{si } A \supset B, \quad \bar{A} \subset \bar{B}$$

(cette relation est d'une importance toute particulière!) (voir fig. 21). Ainsi, du fait que l'ensemble des bons élèves est plus grand que celui des excellents élèves, il s'ensuit que l'ensemble des élèves faibles est contenu dans l'ensemble des élèves qui ne sont pas excellents. Jusqu'ici nous avons insisté tout particulièrement sur l'analogie des relations $\supset$ pour les ensembles et $\geqslant$ pour les nombres. Nous allons signaler maintenant la distinction qui existe entre ces deux relations. Deux nombres (réels) quelconques a et b peuvent être comparés entre eux, c'est-à-dire au moins l'une des deux relations $a \geqslant b$ et $b \geqslant a$ a toujours lieu *). Par contre, pour deux ensembles A, et B quelconques aucune des deux relations $A \supset B$ et $B \supset A$ n'est, en général, vérifiée (voir fig. 22).

Il est également à remarquer que pour tout élément A de l'algèbre des ensembles on a

$$I \supset A, \quad A \supset O,$$

de même qu'on a toujours (quels que soient A et B)

$$A + B \supset A, \ AB \subset A$$

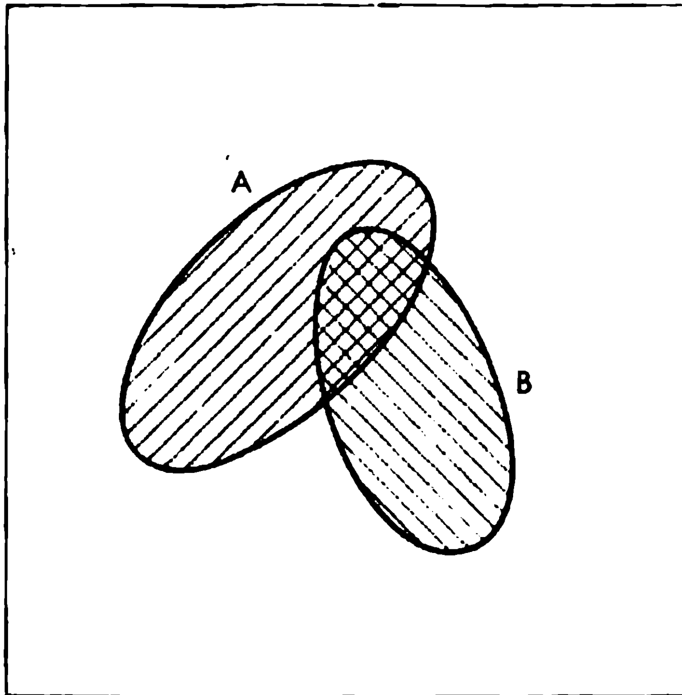
(voir fig. 23).

Elucidons maintenant le sens de la relation $\supset$ pour les autres algèbres de Boole que nous connaissons. Pour l'« algèbre de deux nombres » (exemple 1, p. 31) cette relation est déterminée par la condition suivante

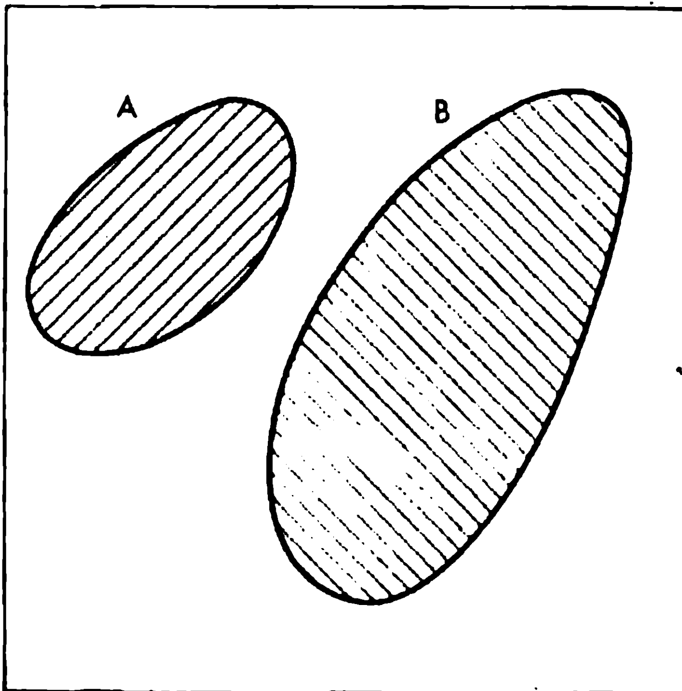
$$1 \supset 0$$

---

*) Si ces deux relations ont simultanément lieu, cela prouve que les nombres a et b sont égaux.



a)



b)

Fig. 22

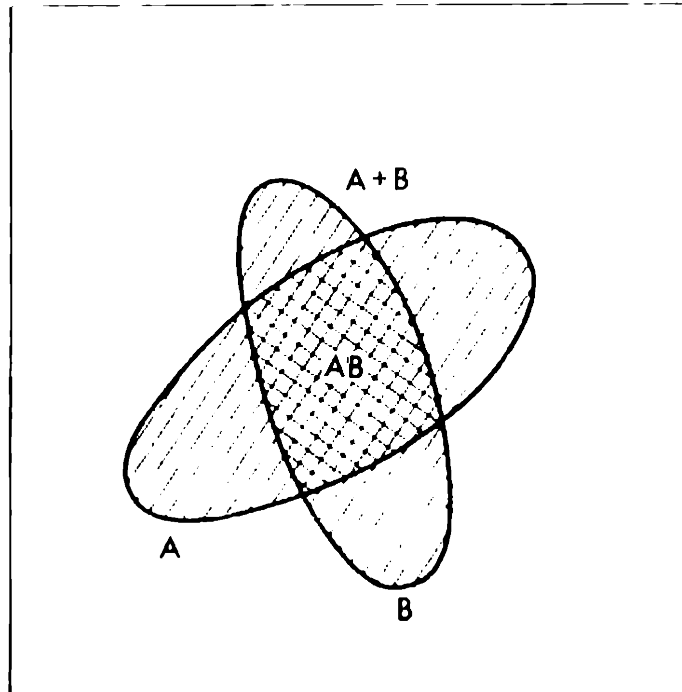


Fig. 23

et pour l'« algèbre de quatre nombres » (exemple 2, p. 32) par les conditions

$$1 \supset 0,\ 1 \supset p,\ 1 \supset q,\ p \supset 0 \text{ et } q \supset 0$$

(les éléments p et q de cette algèbre sont *incomparables*, c'est-à-dire aucune des relations $p \supset q$ et $q \supset p$ n'a lieu). Pour l'« algèbre des maxima et des minima » (exemple 3, p. 33) la relation $\supset$ coïncide avec la relation $\geqslant$: nous considérons que les éléments x et y de cette algèbre sont liés par la relation $x \supset y$ si le nombre x *n'est pas inférieur* au nombre y (dans ce cas, par exemple, $1/2 \supset 1/3$ et $1 \supset 1$) *). Enfin, dans l'« algèbre des plus petits multiples et des plus grands diviseurs » (exemple 4, p. 36) la relation $m \supset n$ signifie, enfin, que le nombre n est un *diviseur* du nombre m ; on a par exemple ici $42 \supset 6$, tandis que les nombres 42 et 35 sont incomparables dans cette algèbre (c'est-à-dire aucune des relations $42 \supset 35$ et $42 \subset 35$ n'a lieu). Nous laissons au lecteur le soin de vérifier que la relation $\supset$ ainsi définie possède dans chacune des algèbres de Boole examinées ci-dessus toutes les propriétés mentionnées de la relation $\supset$ dans l'algèbre des ensembles.

*) Dans cette algèbre de Boole deux éléments *quelconques* x et y sont toujours liés par l'une au moins des relations $x \supset y$ et $y \supset x$.

Il est donc naturel d'appeler *inégalité booléenne* une formule dont les deux membres sont liés par la relation $\supset$ (ou $\subset$). Nous n'examinerons dans ce qui suit que les inégalités qui sont vérifiées pour toutes les valeurs des éléments A, B, C, . . . de l'algèbre de Boole intervenant dans cette inégalité, comme par exemple, $I \supset A$, $A \supset O$, $A + B \supset A$ ou $A \supset AB$. Le *principe de dualité* affirme qu'*en remplaçant dans une telle inégalité l'addition par la multiplication et inversement*, *l'élément* O (s'il faisait partie de cette inégalité) *par* I *et inversement et le signe de l'inégalité par le signe inverse* (c'est-à-dire en remplaçant la relation $\supset$ par $\subset$), *on obtient de nouveau une inégalité vraie* (c'est-à-dire celle qui est vérifiée pour toutes les valeurs des éléments de l'algèbre de Boole qui y figurent). Ainsi, la relation

$$(A + B)(A + C)(A + I) \supset ABC$$

(voir exercice 8,b, p. 60) entraîne toujours

$$AB + AC + AO \subset A + B + C.$$

Pour démontrer le principe de dualité, il suffit d'appliquer l'opération « barre » aux deux membres de l'inégalité initiale. Ainsi, l'inégalité $(A + B)(A + C)(A + I) \supset ABC$, compte tenu de la règle: « si $A \supset B$, alors $\overline{A} \subset \overline{B}$ », entraîne l'inégalité suivante

$$\overline{(A+B)(A+C)(A+I)} \subset \overline{ABC}$$

est toujours vérifiée.

En vertu des règles de De Morgan et compte tenu de ce que $\overline{I} = O$, on obtient

$$\overline{(A+B)(A+C)(A+I)} = \overline{(A+B)(A+C)} + \\ + \overline{A+I} = \overline{A+B} + \overline{A+C} + \overline{AI} = \overline{A}\,\overline{B} + \overline{A}\,\overline{C} + \overline{A}\,\overline{O}.$$

D'une façon analogue

$$\overline{ABC} = \overline{A} + \overline{B} + \overline{C}.$$

Nous pouvons donc en déduire que quels que soient A, B et C l'inégalité suivante

$$\overline{A}\,\overline{B} + \overline{A}\,\overline{C} + \overline{A}O \subset \overline{A} + \overline{B} + \overline{C}$$

a lieu.

Mais puisque dans ce cas $\bar{A}$, $\bar{B}$, $\bar{C}$ sont également arbitraires, on peut donc les désigner tout simplement par A, B et C. On est ainsi amené à l'inégalité

$$AB + AC + AO \subset A + B + C$$

duale de l'inégalité initiale dans le sens indiqué plus haut.

Exercices

1. Trouver les égalités duales de toutes les égalités dont la démonstration fait l'objet des exercices 1-10 à la page 27-28.

2. Démontrer les identités suivantes de l'algèbre des ensembles:

a) $(A + B)(A + \bar{B}) = A$;

b) $AB + (A + B)(\bar{A} + \bar{B}) = A + B$;

c) $\overline{\bar{A}\bar{B}C}\,\overline{\bar{A}B}\,\bar{A}C = O$;

d)* $A + \bar{A}B = A + B$.

3. Démontrer que si l'opération « barre » intervient dans une égalité quelconque de l'algèbre de Boole, on aura une égalité dans laquelle l'addition booléenne est remplacée par la multiplication booléenne et inversement, l'élément O (s'il figure dans l'égalité considérée) par l'élément I et inversement, mais l'opération « barre » occupe la même place que dans l'égalité initiale. [*Exemple*: il suit de l'identité de l'exercice 2,c que

$$\overline{\bar{A}+\bar{B}+C}+\overline{\bar{A}+\bar{B}}+\bar{A}+C=I,$$

quels que soient les éléments A, B, C de l'algèbre de Boole].

4. Appliquer le principe de dualité énoncé dans l'exercice 3 aux égalités de l'exercice 2, a, b, d pour en déduire de nouvelles égalités.

5. Vérifier que la deuxième règle de De Morgan: $\overline{ab} = \bar{a} + \bar{b}$ s'applique également à l'« algèbre de quatre nombres » (exemple 2, p. 32).

6*. a) Soit $N = p_1 p_2 \ldots p_k$, où tous les facteurs premiers sont différents. Démontrer que l'« algèbre des plus petits multiples et des plus grands diviseurs » dont les éléments sont les diviseurs du nombre N (voir exemple 4, p. 36) se ramène dans ce cas à l'« algèbre des sous-ensembles de l'ensemble universel $I = \{p_1, p_2, \ldots, p_k\}$ »; en déduire

que dans cette « algèbre des plus petits multiples et des plus grands diviseurs » toutes les lois de l'algèbre de Boole, les règles de De Morgan comprises, sont vérifiées.

b) Soit $N = p^A$, où p est un nombre premier, et A un entier positif. Démontrer que dans ce cas l'« algèbre des plus petits multiples et des plus grands diviseurs », dont les éléments sont les diviseurs du nombre N, se ramène à l'« algèbre des maxima et des minima » définie dans l'ensemble des nombres 0, 1, 2, . . ., A. En déduire que cette « algèbre des plus petits multiples et des plus grands diviseurs » obéit à toutes les lois de l'algèbre de Boole, les règles de De Morgan comprises.

c) Soient $N = p_1^{A_1} p_2^{A_2} \ldots p_k^{A_k}$, $m = p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}$, où $0 \leqslant a_1 \leqslant A_1$, $0 \leqslant a_2 \leqslant A_2$, . . ., $0 \leqslant a_k \leqslant A_k$ (cf. exercice 6, p. 41). Décomposer le nombre $\overline{m} = \frac{N}{m}$ en facteurs premiers. Démontrer, en faisant appel à la formule ainsi obtenue, les règles de De Morgan pour l'« algèbre générale des plus petits multiples et des plus grands diviseurs ».

7*. Dans quelles algèbres de Boole que nous connaissons, les égalités

$$A + \overline{A} = I \quad \text{et} \quad A\overline{A} = O$$

sont-elles vérifiées et dans quelles algèbres ne le sont-elles pas ?

8. Démontrer les inégalités suivantes de l'algèbre des ensembles:

a) $A + B + C \supset (A + B)(A + C)$;

b) $(A + B)(A + C)(A + I) \supset ABC$;

c) $(A + B)(B + C)(C + A) \supset ABC$;

d) $A + B \supset \overline{A}B + A\overline{B}$.

9. Ecrire les inégalités déduites des inégalités de l'exercice 8, a-c suivant le principe de dualité; démontrer-les directement sans faire appel à ce principe.

10. Démontrer que si dans une inégalité booléenne quelconque figure l'opération « barre » alors sera également vraie l'inégalité déduite de la première en remplaçant l'addition booléenne par la multiplication booléenne et inversement, l'élément O par I et inversement, l'opération « barre » occupant sa place initiale et le signe de l'inégalité se changeant en son inverse. Utiliser ce principe pour obtenir une nouvelle inégalité à partir de l'inégalité de l'exercice 8,d.

11. Vérifier toutes les propriétés de la relation $\supset$ pour:

a) l'« algèbre des maxima et des minima »;

b) l'« algèbre des plus petits multiples et des plus grands diviseurs ».

12*. Soient A et B deux ensembles tels que $A \supset B$. Simplifier les expressions suivantes:

a) $A + B$; b) AB; c) $A + \overline{B}$; d) $\overline{A}B$.

§ 4. ENSEMBLES ET PROPOSITIONS; ALGÈBRE DES PROPOSITIONS

Revenons à l'algèbre booléenne des ensembles. La question se pose donc de savoir comment peut-on se donner des ensembles qui soient éléments de cette algèbre. Il est évident que le procédé le plus simple est de se les donner sous une forme *explicite*, c'est-à-dire tout simplement de les *énumérer*; ainsi, on peut parler de l'« ensemble des élèves: Jean, Paul, Louis, Marie », de l'« ensemble des nombres: 1, 2, 3, 4, 5 » ou de l'« ensemble des quatre opérations arithmétiques: l'addition, la soustraction, la multiplication et la division ». En énumérant les éléments d'un ensemble quelconque, il est d'usage de les mettre entre accolades; ainsi, on peut écrire

$A = \{$Jean, Paul, Louis, Marie$\}$;

$$B = \{1, 2, 3, 4, 5\} \text{ ou } C = \{+, -, \times, :\}$$

(dans cette dernière formule les signes symbolisent les opérations elles-mêmes) *).

Cependant un tel procédé de se donner des ensembles est particulièrement incommode si le nombre d'éléments est élevé; de plus, il est absolument inutilisable pour la donnée des ensembles *infinis*. Par ailleurs, même s'il s'avère possible de se donner explicitement des ensembles, cela embrouille parfois le sens même de l'ensemble considéré, les raisons qui

*) Voir également exercice 6,a, p. 59.

nous ont incité à réunir dans tel ou tel ensemble des éléments bien déterminés et non pas d'autres.

Il existe un autre procédé de se donner des ensembles : c'est le procédé *descriptif* ou *implicite* qui est plus répandu. Dans ce cas on désigne la *propriété* qui est commune à tous les éléments de l'ensemble considéré : ainsi, on peut parler de l'« ensemble de tous les excellents élèves d'une classe quelconque » (il est possible que ce soit précisément l'ensemble A mentionné plus haut) ou de l'« ensemble de tous les nombres entiers x tels que $0 < x \leqslant 5$ » (c'est exactement l'ensemble B) ou, enfin, de l'« ensemble de toutes les bêtes du zoo de Moscou ». Il est à noter que le procédé descriptif peut bien servir aussi pour prélever des ensembles infinis tels que l'« ensemble de tous les nombres entiers » ou l'« ensemble de tous les triangles d'aire 1 » ; bien plus, comme nous l'avons souligné plus haut, les ensembles infinis *ne* peuvent être donnés *que* par le procédé descriptif.

Le procédé implicite (descriptif) lie les ensembles aux *propositions* qu'on étudie en logique mathématique. Ce procédé est le suivant : on fixe un certain ensemble des objets qui seuls nous intéressent (l'ensemble des élèves d'une classe quelconque ou l'ensemble des nombres entiers, par exemple), ensuite on énonce une proposition quelconque à laquelle satisfont tous les éléments de l'ensemble considéré et eux seuls ; si ce sont les ensembles des élèves d'une classe quelconque qui nous intéressent, ces propositions peuvent être : « c'est un excellent élève », « il sait jouer aux échecs », « sa place est au premier rang », « il s'appelle André », etc. L'ensemble A de tous les éléments de l'ensemble universel I considéré (l'ensemble I peut être l'ensemble des élèves, des nombres, etc.) qui satisfont à la propriété d'une proposition donnée a s'appelle *ensemble de vérité* de cette proposition (voir, par exemple, fig. 24) *).

*) Nous désignerons toujours les propositions par des lettres minuscules et les ensembles de vérité qui leur correspondent de préférence par les *mêmes lettres* majuscules.

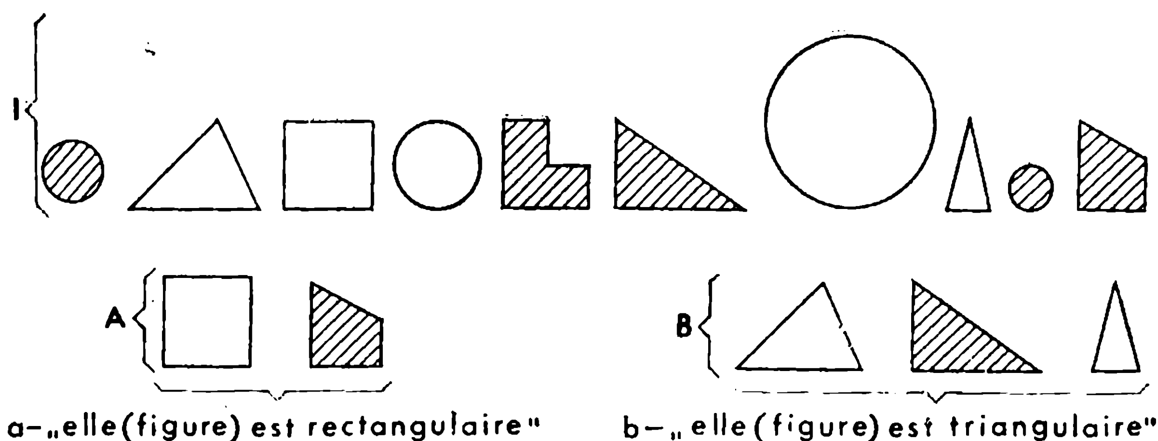


Fig. 24

De cette façon, nous avons établi une « liaison bilatérale » entre les ensembles et les propositions : tout ensemble est décrit par une certaine proposition (celle-ci peut comporter une simple énumération des éléments d'un ensemble : « c'est Jean, ou Paul, ou Louis, ou André ») et à chaque proposition correspond un ensemble de vérité bien déterminé de cette proposition. De plus, pour tout ensemble des propositions — même celles ayant trait aux objets les plus divers — on peut toujours indiquer un ensemble universel I contenant *tous* les objets dont il s'agit dans ces propositions. Néanmoins, et ceci est d'ailleurs très important, *par une proposition nous entendrons une affirmation telle qu'on puisse dire si elle est vraie* (relativement à un élément déterminé de l'ensemble universel considéré) *ou fausse*. Ainsi, les phrases telles que « il a deux têtes et seize bras » ou « $2 \times 3 = 6$ » sont des propositions (la deuxième phrase ne dépend pas même du tout du choix de l'ensemble universel I), tandis que le slogan « Vive la Paix ! » ou l'interjection « Oh ! » ne le sont évidemment pas.

Puisque ces propositions ne nous intéressent que du point de vue des ensembles qu'elles décrivent, nous ne ferons pas de distinction entre deux propositions a et b auxquelles correspond le même ensemble de vérité, tout en les

considérant comme identiques. Si les deux propositions a et b (« c'est un excellent élève » et « il n'a que d'excellentes notes » ou « le nombre x est impair » et « le nombre x divisé par 2 donne 1 pour reste ») sont identiques, on écrira

$$a = b.$$

De plus, nous considérerons comme identiques toutes les propositions *identiquement vraies* (ou *vides de sens*), c'est-à-dire des propositions qui sont *toujours* vraies quel que soit l'élément de l'ensemble I considéré: ainsi, les propositions « $2 \times 3 = 6$ », « il mesure moins de 3 m », etc., sont identiquement vraies. Nous conviendrons de désigner toutes les propositions vraies par i. De façon analogue, nous considérerons comme étant identiques toutes les propositions *identiquement fausses* (ou *contradictoires*) qui n'ont *jamais* lieu, c'est-à-dire des propositions dont l'ensemble de vérité est vide. Citons quelques exemples de telles propositions que nous désignerons par o: « $2 \times 2 = 6$ », « il (un élève, par exemple) sait voler », « il mesure plus de 4 m », « ce nombre est plus grand que 3 mais plus petit que 2 ».

Les liens existant entre les ensembles et les propositions permettent de définir des *opérations algébriques* particulières, proches des opérations de l'algèbre des ensembles introduites plus haut. Notamment, *nous appellerons somme de deux propositions a et b une proposition dont l'ensemble de vérité coïncide avec la somme de l'ensemble de vérité A de la proposition a et de l'ensemble de vérité B de la proposition b. Nous conviendrons de noter cette proposition $a + b$* *). Or la somme de deux ensembles n'est autre que la réunion de tous les éléments que ces deux ensembles contiennent, donc la somme de deux propositions a et b est la proposition « a ou b »,

---

*) En logique mathématique la somme de deux propositions a et b est généralement appelée *disjonction* de ces propositions et notée $a \vee b$ (cf. avec la notation $A \cup B$ d'une somme de deux ensembles A et B).

« ou » signifiant que soit la proposition a, soit la proposition b, soit, enfin, ces deux propositions à la fois, sont vraies. Si par exemple la proposition a est la suivante : « il sait jouer aux échecs » et dans une classe quelconque à cette proposition correspond l'ensemble de vérité

$A =$ {Jean, Paul, Louis, André, Marie, Edith, Anne},

et la proposition b : « il sait jouer aux dames » et admet pour ensemble de vérité

$B =$ {Jean, Louis, Guy, Ives, Marie, Ella},

alors $a + b$ représente la proposition « il sait jouer aux échecs *ou* il sait jouer aux dames » (ou tout simplement « il sait jouer aux échecs *ou* aux dames ») à laquelle correspond l'ensemble de vérité

$A + B =$ {Jean, Paul, Louis, André, Guy, Ives, Marie, Edith, Anne, Ella}.

Si l'ensemble universel est l' ensemble des figures représentées à la figure 24 et si les propositions c et d ont un sens : « elle (la figure) est ronde » et « elle (la figure) est hachurée », alors la proposition $c + d$ s'énonce ainsi : « elle (la figure) est ronde *ou* hachurée » (voir fig. 25).

D'une façon analogue, on appelle *produit ab de deux propositions a et b dont les ensembles de vérité sont respectivement A et B une proposition dont l'ensemble de vérité coïncide avec le produit AB des ensembles A et B* *). Or, le produit de deux ensembles A et B est leur *intersection* ou leur *partie commune* qui ne contient que les éléments appartenant simultanément à A et B ; pour cette raison, le produit ab de deux propositions a et b est une proposition « a et b », « et »

*) En logique mathématique le produit de deux propositions a et b est plus souvent appelé *conjonction* et noté $a \wedge b$ (cf. avec la notation $A \cap B$ d'un produit de deux ensembles A et B).

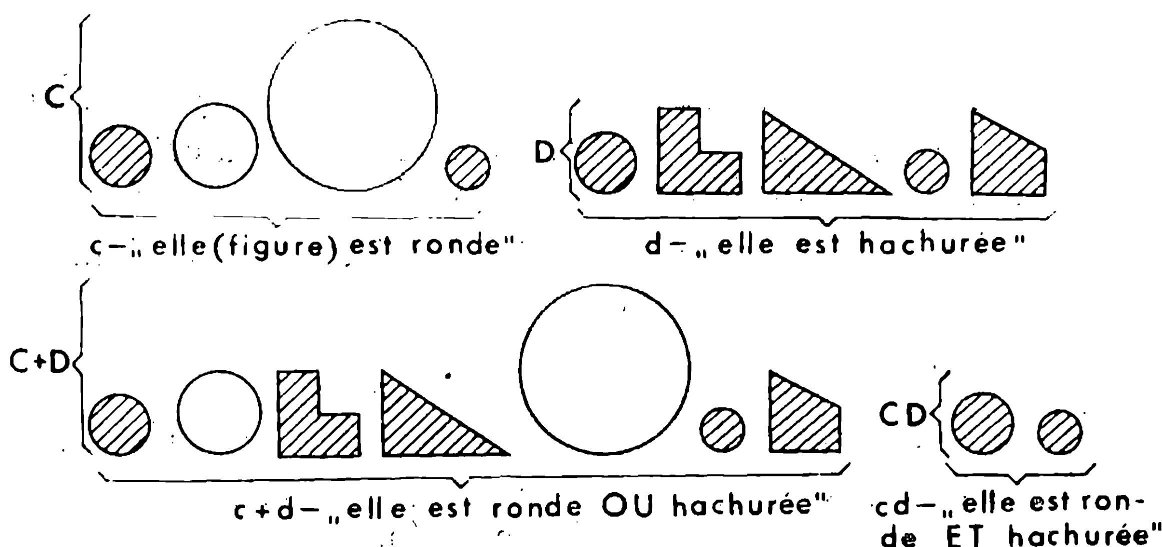


Fig. 25

signifiant, comme toujours, que *les deux propositions* sont vraies: aussi bien a que b. Si, par exemple, les propositions a et b ont le même sens que plus haut, la proposition ab signifie: « il sait jouer aux échecs *et* il sait jouer aux dames » (ou pour abréger, « il sait jouer aux échecs *et* aux dames ») et à cette proposition correspond l'ensemble de vérité

$$AB = \{\text{Jean, Louis, Marie}\}.$$

Si les propositions c et d qui se rapportent à l'ensemble des figures représenteés sur la figure 24, ont un sens: « elle (la figure) est ronde » et « elle (la figure) est hachurée », alors la proposition cd signifie que « cette figure est ronde *et* hachurée » (voir fig. 25).

La liaison qui existe entre les ensembles et les propositions permet d'appliquer aux propositions toutes les règles de l'algèbre des ensembles:

$$a + b = b + a, \qquad ab = ba;$$

commutativité de l'algèbre des propositions

$$(a + b) + c = a + (b + c), \quad (ab)\,c = a\,(bc);$$

associativité de l'algèbre des propositions

$$(a+b)\,c = ac + bc, \quad ab + c = (a+c)\,(b+c)\,;$$

distributivité de l'algèbre des propositions

$$a + a = a, \qquad aa = a.$$

idempotence de l'algèbre des propositions

Si, de plus, i est une proposition identiquement vraie et o une proposition identiquement fausse, alors on a toujours (c'est-à-dire *quelle que soit* la proposition a)

$$a + o = a, \quad ai = a\,;$$
$$a + i = i, \quad ao = o.$$

Ainsi, la proposition « c'est un excellent élève ou il a deux têtes » équivaut à la proposition « c'est un excellent élève », et la proposition « il sait nager et il a moins de 200 ans » équivaut à la proposition « il sait nager » *).

Pour mieux comprendre comment les règles de l'algèbre des propositions se déduisent des règles de l'algèbre des ensembles, examinons, par exemple, la *deuxième loi de distributivité*. Etant donné que l'ensemble de vérité d'une somme de deux propositions est la somme des ensembles de vérité de ces propositions et l'ensemble de vérité d'un produit de propositions est le produit de leurs ensembles de vérité, l'ensemble de vérité d'une proposition composée $ab + c$, c'est-à-dire de la proposition « on a « *a et b* » *ou c* » est $AB + C$, où A, B et C sont respectivement les ensembles de vérité des propositions a, b et c. D'une façon analogue, l'ensemble de vérité de la proposition (composée) $(a + c)\,(b + c)$ est l'ensemble $(A + C)\,(B + C)$. Mais, en vertu

---

*) Ecrivons maintenant les règles ci-dessus énoncées sous la forme admise en logique mathématique:

$$a \vee b = b \vee a, \qquad a \wedge b = b \wedge a,$$
$$(a \vee b) \vee c = a \vee (b \vee c), \qquad (a \wedge b) \wedge c = a \wedge (b \wedge c),$$
$$(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c), \qquad (a \wedge b) \vee c = (a \vee c) \wedge (b \vee c),$$
$$a \vee a = a, \qquad a \wedge a = a,$$
$$a \vee o = a, \qquad a \wedge i = a,$$
$$a \vee i = i, \qquad a \wedge o = o.$$

de la deuxième loi de distributivité de l'algèbre des ensembles, on a

$$AB + C = (A + C)(B + C).$$

Ainsi, les ensembles de vérité des propositions $ab + c$ et $(a + c)(b+c)$ coïncident, ce qui signifie tout simplement que les propositions $ab + c$ et $(a + c)(b + c)$ sont identiques! [Voir également p. 24 où nous avons noté que les propositions « il sait jouer aux échecs et aux dames ou il sait nager » et « il sait jouer aux échecs ou il sait nager, et de plus il sait jouer aux dames ou il sait nager » ont le même sens, c'est-à-dire que

$$ab + c = (a + c)(b + c),$$

où les propositions a, b et c signifient: « il sait jouer aux échecs », « il sait jouer aux dames » et « il sait nager ».]

Outre les opérations d'addition et de multiplication des ensembles, l'opération « barre » peut également s'employer dans l'algèbre des propositions. Par $\bar{a}$ nous entendrons alors la *proposition dont l'ensemble de vérité est l'ensemble* $\bar{A}$, A *étant l'ensemble de vérité de la proposition* a. En d'autres termes, la condition $\bar{a}$ doit être vérifiée seulement par les éléments de l'ensemble universel I qui *ne figurent pas* dans l'ensemble A, c'est-à-dire par les éléments qui *ne vérifient pas* la condition a. Soit la proposition a: « il a de mauvaises notes », la proposition $\bar{a}$ signifiera: « il n'a pas de mauvaises notes » (« il a de bonnes notes dans toutes les disciplines »); si l'ensemble universel I se compose des figures représentées sur la figure 24 et la proposition b signifie: « elle (la figure) est triangulaire », la proposition $\bar{b}$ aura le sens: « elle (la figure) n'est pas triangulaire » (fig. 26). D'une façon générale, la proposition $\bar{a}$ signifie « **non** a »; il est donc naturel que l'opération « barre » de l'algèbre des propositions se nomme *négation*.

Voici la liste des règles liées à l'opération de négation:

$$\bar{\bar{a}} = a;$$

$$a + \bar{a} = i \quad \text{et} \quad a\bar{a} = o;$$

$$\bar{o} = i \quad \text{et} \quad \bar{i} = o;$$

$$\overline{a + b} = \bar{a}\,\bar{b} \quad \text{et} \quad \overline{ab} = \bar{a} + \bar{b}.$$

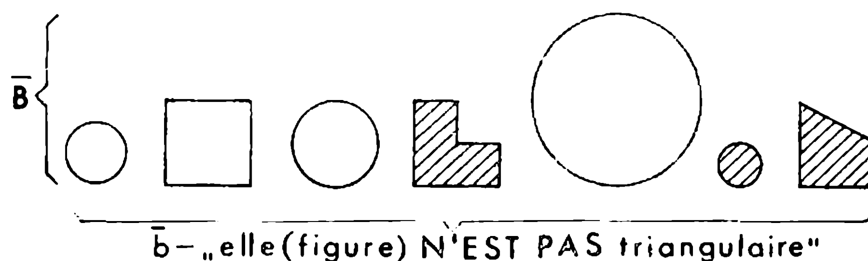


Fig. 26

En effet, la négation d'une proposition identiquement fausse (« 2×2 *n*'est *pas* égal à 5 » ou « cet élève *n*'a *pas* deux têtes », par exemple) est toujours une proposition identiquement vraie et la négation d'une proposition identiquement vraie (« cet élève a plus de 120 ans ») est toujours identiquement fausse. Il est aisé de vérifier la validité de toutes les autres lois (nous laissons ce soin au lecteur) ; elles n'exigent d'ailleurs aucune vérification spéciale puisqu'elles découlent des règles correspondantes de l'algèbre des ensembles *).

Exercices

1. Donner trois exemples de propositions identiquement vraies, deux exemples de propositions identiquement fausses.

2. Soit la proposition a :

a) « $2 \times 2 = 4$ » ;
b) « c'est un garçon » ;
c) « l'éléphant est un insecte » ;
d) « il sait voler ».

Quel est, dans tous ces cas, le sens de la proposition $\bar{a}$? Est-elle une proposition identiquement vraie? Est-elle identiquement fausse?

3. Soient la proposition a : « il sait jouer aux échecs » et la proposition b : « il sait jouer aux dames ». Définir les propositions suivantes :

a) $a+b$; b) ab ; c) $\bar{a}+b$; d) $a+\bar{b}$; c) $\bar{a}+\bar{b}$; f) $\bar{a}b$; g) $a\bar{b}$; h) $\bar{a}\bar{b}$.

*) Ainsi, vu que les ensembles de vérité des propositions $\overline{a+b}$ et $\bar{a}\bar{b}$ sont égaux à $\overline{A+B}$ et $\bar{A}\bar{B}$ respectivement, où A et B sont les ensembles de vérité des propositions a et b, et $\overline{A+B} = \bar{A}\bar{B}$, alors, d'après la définition de l'égalité (coïncidence) des propositions, on a $\overline{a+b} = \bar{a}\bar{b}$.

4. Soient les propositions a: « c'est un excellent élève », b: « il est brun », c: « il sait nager ». Définir les propositions suivantes:

a) $(a+b)\,c$ et $ac+bc$;

b) $ab+c$ et $(a+c)\,(b+c)$.

5. Soient les deux propositions a et b: « c'est un nombre (entier positif) pair » et « c'est un nombre premier ». Définir les propositions:

a) ab; b) $\bar{a}+b$; c) $\bar{a}b$; d) $a\bar{b}$; c) $\bar{a}+\bar{b}$.

Quels sont les ensembles de vérité de ces propositions?

6. Soient les propositions a et b: « il fréquente le cercle de mathématiques » et « c'est un choriste ». Définir les propositions suivantes:

a) $\overline{a+b}$ et $\bar{a}\,\bar{b}$;

b) $\overline{ab}$ et $\bar{a}+\bar{b}$.

§ 5. « LOIS DE LA PENSÉE » ET RÈGLES DE DÉDUCTION

Maintenant nous sommes en mesure de dire pourquoi G. Boole a intitulé son ouvrage, où sont exposés les principes de l'« algèbre non ordinaire » que nous examinons ici, « L'étude des lois de la pensée ». En effet, l'*algèbre des propositions* est intimement liée aux règles qui guident la pensée humaine, car après tout la somme et le produit de propositions définis plus haut ne signifient rien d'autre que les liaisons logiques « ou » et « et », l'opération « barre » étant une négation tandis que les lois de l'algèbre des propositions décrivent les propriétés fondamentales de ces *opérations logiques* effectuées par la pensée humaine. Certes, rares sont ceux qui perçoivent ces propriétés comme les lois mathématiques de la pensée, bien que même les petits enfants *s'en servent* déjà facilement. En effet, personne ne contestera que: « courir vite et sauter haut » est la même chose que: « sauter haut et courir vite »; en d'autres termes, tout le monde sait (consciemment ou non) que les propositions ab et ba ont le même sens, c'est-à-dire sont « égales ».

On peut maintenant se faire une idée du regain d'intérêt suscité aujourd'hui par les recherches de G. Boole et par l'interprétation mathématique des lois logiques sous forme de « règles algébriques » originales. Tant que la pensée n'a été que la prérogative absolue de la raison humaine on ne s'est guère soucié de formaliser les « lois de la pensée »: c'est que les hommes se sont toujours laissés guider par ces lois sans en saisir pleinement le contenu. Ces dernières décennies ont littéralement bouleversé cette situation, aujourd'hui nous aspirons toujours plus à faire assumer aux calculateurs électroniques (nos « aides électroniques ») des fonctions qui, tout récemment encore, étaient l'apanage de la seule raison humaine: gestion de l'économie, réglementation des transports, résolution de problèmes mathématiques, traduction de livres, planification de l'économie, recherche de l'information nécessaire; aujourd'hui enfin, les calculateurs savent même jouer aux échecs! Mais pour apprendre tout cela aux calculateurs, il nous a fallu naturellement formuler d'une façon rigoureuse ces « règles du jeu », ces « lois de la pensée » auxquelles doivent se conformer les appareils « intelligents » construits par l'homme : si l'homme se conforme *instinctivement* aux lois de la logique, pour les calculateurs par contre, ces lois doivent être formulées de façon précise dans le « langage » mathématique, le seul que ces machines sont susceptibles de « comprendre », celui des mathématiques *).

Revenons aux « lois de la pensée ». Les plus intéressantes sont les règles qui sont liées à l'opération logique de négation. Nombre d'entre elles sont explicitement désignées.

---

*) A ne pas en conclure que l'algèbre élémentaire des propositions qui est traitée dans ce livre, représente à elle seule un appareil mathématique suffisant pour construire des calculateurs complexes ou bien pour poser des problèmes sous une forme telle que leur résolution pourra être « confiée » aux appareils électroniques.

Ainsi, la règle

$$a + \bar{a} = i$$

est appelée *principe du tiers exclu* : ou bien a lieu la proposition a, ou bien la proposition $\bar{a}$, et pas une tierce, donc la proposition $a + \bar{a}$, c'est-à-dire « a ou non a » est toujours vraie. Ainsi, sans rien savoir du « plus grand élève de septième de l'école n° 12 de Léningrad » nous pouvons affirmer: « ou bien c'est un excellent élève ou bien ce n'est pas un excellent élève », ou encore « il sait jouer aux échecs ou il ne sait pas jouer aux échecs ». La règle

$$a\bar{a} = o$$

est appelée *principe de contradiction*; elle affirme que les propositions a et $\bar{a}$, c'est-à-dire a et « non a » ne peuvent jamais avoir lieu simultanément, autrement dit leur produit est toujours faux. Si par exemple un élève quelconque est réellement un excellent élève, alors la proposition « ce n'est pas un excellent élève » est certainement fausse ; si le nombre (entier) n est pair, la proposition « n est impair » est évidemment fausse. La règle

$$\bar{\bar{a}} = a$$

est appelée *loi de double négation*; elle signifie que la double négation d'une affirmation quelconque équivaut à l'affirmation initiale. Ainsi, la négation de la proposition « n est pair » est la proposition « n est impair »; la nouvelle négation « n n'est pas impair » de cette dernière proposition nous amène à la proposition initiale, à savoir: « n est pair ». D'une façon analogue, la double négation « ce n'est pas un élève faible » de l'affirmation que c'est un bon élève équivaut à la proposition initiale « c'est un bon élève ».

Pour les propositions dont la formulation verbale est plus compliquée (voir l'exercice 1 ci-dessous) les règles de

De Morgan

$$\overline{a+b}=\bar{a}\,\bar{b} \quad \text{et} \quad \overline{ab}=\bar{a}+\bar{b}$$

gardent tout leur intérêt. Les autres règles de l'algèbre des propositions, telles que les lois de distribution

$$(a+b)\,c = ac + bc \quad \text{et} \quad ab + c = (a+c)\,(b+c)$$

ou les lois d'idempotence

$$a + a = a \quad \text{et} \quad aa = a,$$

sont également des « lois de la pensée », des règles logiques permettant de déduire de nouvelles conséquences à partir de celles déjà connues.

Dans cet ordre d'idées, une place particulière revient aux règles liées à la relation $\supset$. Nous ne l'avons pas encore examinée; néanmoins la « liaison bilatérale » établie plus haut entre les ensembles et les propositions permet d'appliquer facilement la relation $\supset$ de l'algèbre des ensembles (c'est-à-dire la relation d'inclusion) à celle des propositions. Ainsi, nous écrirons

$$a \supset b$$

pour dire que la *proposition a découle de la proposition b* (ou *a est une conséquence de b*), si l'*ensemble de vérité A de la proposition a comprend l'ensemble de vérité B de la proposition b*, c'est-à-dire si

$$A \supset B.$$

Ainsi, du moment que l'ensemble B des excellents élèves d'une classe est, de toute évidence, contenu dans l'ensemble A des meilleurs élèves de la même classe, la proposition a: « cet élève est fort en toutes les disciplines » est une conséquence de la proposition b: « c'est (il s'agit du même élève) un excellent élève ». D'une façon analogue, l'ensemble

$$A_6 = \{6,\ 12,\ 18,\ \ldots\}$$

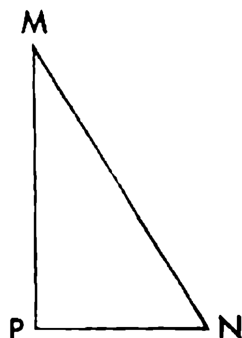


Fig. 27

des multiples (entiers positifs) de 6 est contenu dans l'ensemble,

$$A_2 = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20, \ldots\}$$

des nombres pairs; la proposition « il (un nombre) est pair » est donc une conséquence de la proposition « c'est un multiple de 6 » *).

On appelle souvent *déduction* la relation $a \supset b$ entre les propositions a et b; la proposition b est alors appelée *hypothèse* et la proposition a qui en résulte *conclusion*. On rencontre les déductions dans la science et dans la vie quotidienne: ainsi, toutes les démonstrations des théorèmes mathématiques revêtent en général le caractère d'une déduction: il faut établir que l'hypothèse b d'un théorème (« l'angle P du triangle MNP est droit », par exemple, voir fig. 27) entraîne la conclusion a (« $MP^2 + NP^2 = MN^2$ », dans ce cas la formule $a \supset b$ équivaut au *théorème de Pythagore*). Dans les déductions (pour la démonstration des théorèmes, par exemple) on fait constamment appel (parfois inconsciemment) aux propriétés fondamentales de la rela-

*) Si $a \supset b$, on dit alors que la proposition b est une *condition suffisante* pour a (pour qu'un élève soit fort en toutes les matières il est évidemment *suffisant* qu'il soit un excellent élève) et la proposition a, une *condition nécessaire* pour b (pour qu'un élève soit excellent il est bien sûr *nécessaire* qu'il soit fort en toutes les matières).

tion $\supset$*) :

$$a \supset a;$$

$$\text{si } a \supset b \quad \text{et} \quad b \supset a, \text{ alors } a = b;$$

$$\text{si } a \supset b \quad \text{et} \quad b \supset c, \text{ alors } a \supset c;$$

$$i \supset a \quad \text{et} \quad a \supset o \text{ quel que soit } a;$$

$$a + b \supset a \quad \text{et} \quad a \supset ab \text{ quels que soient } a \text{ et } b;$$

$$\text{si} \quad a \supset b, \text{ alors } \bar{b} \supset \bar{a}.$$

On sait, par exemple, que *si les diagonales d'un quadrilatère se coupent en leurs milieux* (proposition b) *ce quadrilatère est un parallélogramme* (proposition a) **) ; d'autre part, les *angles opposés* d'un parallélogramme *sont égaux* (proposition c). Donc, on a

$$a \supset b \quad \text{et} \quad c \supset a;$$

d'où

$$c \supset b,$$

en d'autres termes: *si les diagonales d'un quadrilatère se coupent en leurs milieux, ses angles opposés sont égaux.*

Attardons-nous un peu plus sur cette règle: *si* $a \supset b$, *alors* $\bar{b} \supset \bar{a}$. Cette règle sert de base aux démonstrations dites *par l'absurde*. Soit à démontrer la relation $a \supset b$: la proposition b entraîne la proposition a. Souvent il est peu facile de démontrer que si *a n'a pas lieu*, b n'aura pas lieu non plus, c'est-à-dire la proposition « non a » (proposition $\bar{a}$) entraîne la proposition « non b » (proposition $\bar{b}$).

En voici un exemple: nous allons démontrer que *si un nombre (entier) n, supérieur à* 3, *est premier* (proposition b),

---

*) La règle: *si* $a \supset b$ *et* $b \supset a$, *alors* $a = b$ peut parfois être énoncée comme suit: si b est une *condition nécessaire et suffisante* pour a, les propositions a et b sont alors équivalentes (c'est-à-dire identiques, égales d'après nos définitions).

**) Dans ce cas on a même $a \supset b$ et $b \supset a$, c'est-à-dire $a = b$.

alors n sera de la forme $6k \pm 1$ (*où k est un entier*), c'est-à-dire divisé par 6 le nombre n donne $+1$ ou -1 pour reste (proposition a). La démonstration directe, sans faire appel à la règle « si $a \supset b$, alors $\bar{b} \supset \bar{a}$ », pose des difficultés; voyons ce que donne la démonstration par l'absurde. Supposons qu'on ait la proposition $\bar{a}$, c'est-à-dire que le nombre n (entier et supérieur à 3) *ne soit pas* de la forme $6k \pm 1$. Vu que tout nombre entier n divisé par 6 donne pour reste soit 0 (s'il est un multiple de 6), soit 1, soit 2, soit 3, soit 4, soit 5 (ou ce qui revient au même -1), la proposition $\bar{a}$ signifie que le *nombre n, divisé par* 6, *donne pour reste soit* 0 (c'est-à-dire qu'il est un multiple de 6), *soit* 2, *soit* 3, *soit* 4. Mais un nombre divisible par 6 *a priori* n'est pas premier; si un nombre entier $n > 3$, divisé par 6, donne 2 ou 4 pour reste, il est alors pair d'où il s'ensuit qu'il ne peut pas être premier; si, enfin, n a 3 pour reste, il se divise alors par 3, donc il n'est non plus premier. Ainsi, de $\bar{a}$ découle $\bar{b}$ (ce que l'on note: $\bar{b} \supset \bar{a}$); d'où il s'ensuit immédiatement que

$$a \supset b,$$

ce qu'il fallait démontrer *).

Exercices]

1. Enoncer les règles de De Morgan pour l'algèbre des propositions: $\overline{a+b} = \bar{a}\bar{b}$ et $\overline{ab} = \bar{a} + \bar{b}$.

2. Donner des exemples illustrant le:
   a) principe du tiers exclu;
   b) principe de contradiction;
   c) principe de double négation.

---

*) Le raisonnement suivant est plus rigoureux: la relation démontrée $\bar{b} \supset \bar{a}$ entraîne $\overline{(\bar{a})} \supset \overline{(\bar{b})}$ ou $\bar{\bar{a}} \supset \bar{\bar{b}}$; étant donné qu'en vertu de la loi de double négation $\bar{\bar{a}} = a$ et $\bar{\bar{b}} = b$, on a $a \supset b$.

3. Donner des exemples illustrant chacune des propriétés de la relation $\supset$ énumérées à la page 75.

4. Ecrire sous forme de symboles la démonstration d'un théorème quelconque par l'absurde.

5. Soit $a \supset b$. Simplifier la somme $a + b$ des propositions a et b de même que leur produit ab.

§ 6. PROPOSITIONS ET CHAÎNES DE CONTACTS

Pour conclure, donnons encore un exemple d'algèbre booléenne tout étrange qu'il puisse vous paraître. Pour éléments de notre algèbre nous prendrons toutes les *chaînes de contacts* possibles, c'est-à-dire des circuits électriques ouverts par une série d'interrupteurs (fig. 28). Nous désignerons par des majuscules les diverses branches d'un circuit.

Vu que la seule fonction d'un tronçon de circuit électrique est de transmettre le courant, nous n'allons pas distinguer deux conducteurs identiques, c'est-à-dire des conducteurs munis des mêmes interrupteurs et transmettant ou non le courant lorsque ces derniers occupent la même position (« fermé » ou « ouvert »), nous dirons qu'ils sont « égaux ». Convenons enfin d'appeler *somme* $A + B$ des tronçons A et B d'un circuit leur *connexion en parallèle* et *produit* AB leur *connexion en série* (voir fig. 29,*a*, *b* où les tronçons A et B du circuit comportent chacun un seul contact). Il est clair que l'addition et la multiplication ainsi énoncées des tronçons d'un circuit sont *commutatives*:

$$A + B = B + A, \quad AB = BA$$

et *associatives*:

$$(A + B) + C = A + (B + C) \quad (= A + B + C)$$
$$(AB)\,C = A\,(BC) \quad (= ABC)$$

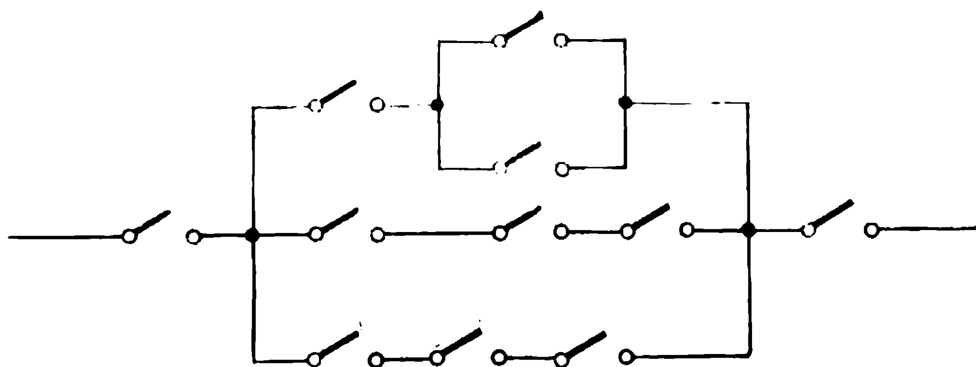


Fig. 28

(voir fig. 30,*a*, *b* où sont représentés la « somme triple » $A + B + C$ des contacts A, B, C et leur « produit triple » ABC). On comprend aisément que lesdites opérations vérifient également les lois *idempotentes*

$$A + A = A, \quad AA = A,$$

car la mise en série ou en parallèle de deux contacts identiques (c'est-à-dire simultanément fermés ou ouverts) donne le même résultat qu'un seul contact. Il est plus difficile de vérifier dans notre « algèbre des chaînes de contacts » la validité des deux lois *distributives*:

$$(A + B)\,C = AC + BC \quad \text{et} \quad AB + C = (A + C)\,(B + C);$$

néanmoins ces lois aussi, comme cela résulte des figures 31 et 32, sont vérifiées ici (il est aisé de voir que le montage représenté sur la figure 31,*a* est « égal », au sens convenu précédemment, à celui de la figure 31,*b*, et le montage de la figure 32,*a* à celui de la figure 32,*b*).

Convenons encore de désigner par I un contact toujours fermé (soudé) (voir fig. 33,*a*) et par O un contact toujours ouvert (rupture de circuit; voir fig. 33,*b*). On a évidemment

$$A + O = A \quad \text{et} \quad AI = A$$

(voir fig. 34);

$$A + I = I \quad \text{et} \quad AO = O$$

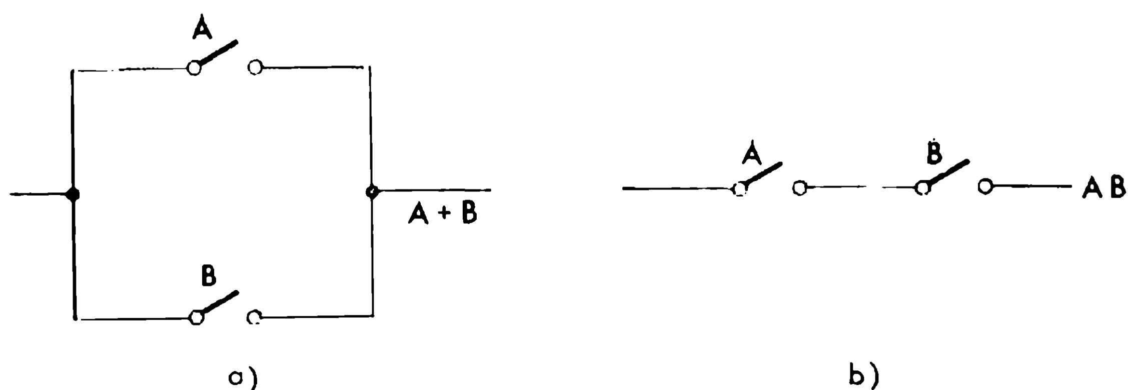


Fig. 29

(voir fig. 35) ; ainsi, le rôle des éléments « particuliers » I et O de notre algèbre de Boole incombe ici aux contacts I et O.

Convenons enfin de désigner par A et $\overline{A}$ un couple de contacts tels que *si le contact A est fermé, le contact $\overline{A}$ est obligatoirement ouvert*; du point de vue technique, la réalisation d'un tel couple de contacts est assez facile (voir fig. 36). Il est évident que

$$\overline{\overline{A}}=A,\ \overline{I}=O \quad \text{et} \quad \overline{O}=I,$$

de même que

$$A+\overline{A}=I \quad \text{et} \quad A\overline{A}=O$$

(voir fig. 37,*a*, *b*). Les règles de De Morgan sont plus difficiles à démontrer:

$$\overline{A+B}=\overline{A}\,\overline{B} \quad \text{et} \quad \overline{AB}=\overline{A}+\overline{B},$$

bien qu'elles soient également satisfaites (voir fig. 38,*a*, *b* où, disons, les *tronçons* $A+B$ et $\overline{A+B}$ du circuit sont définis par la condition suivante: si le courant passe par le tronçon $A+B$, il ne passe pas par le tronçon $\overline{A+B}$ et inversement).

La ressemblance de l'« algèbre des chaînes de contacts » avec l'« algèbre des propositions » est importante sous deux

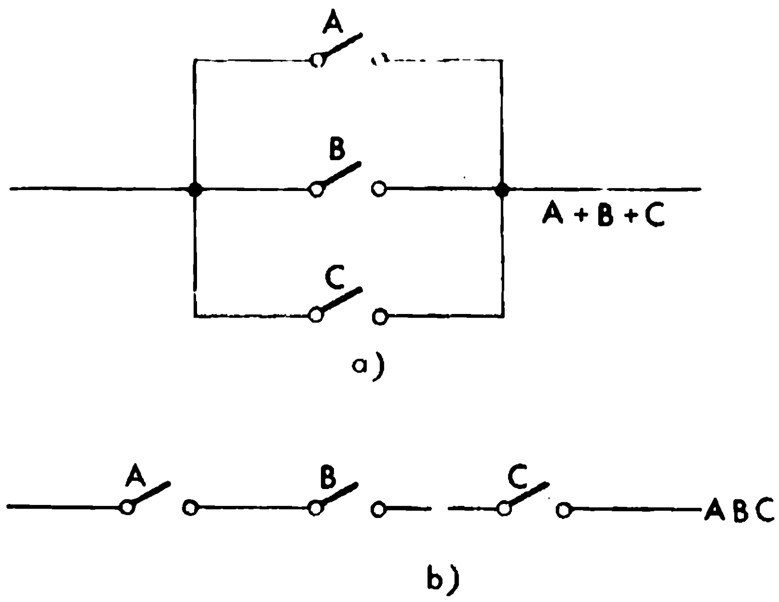


Fig. 30

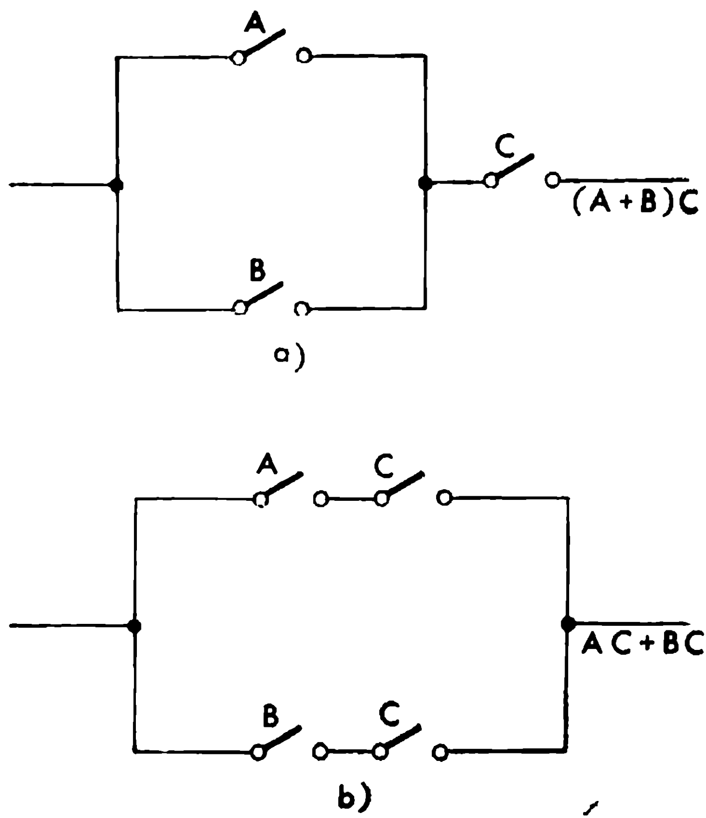
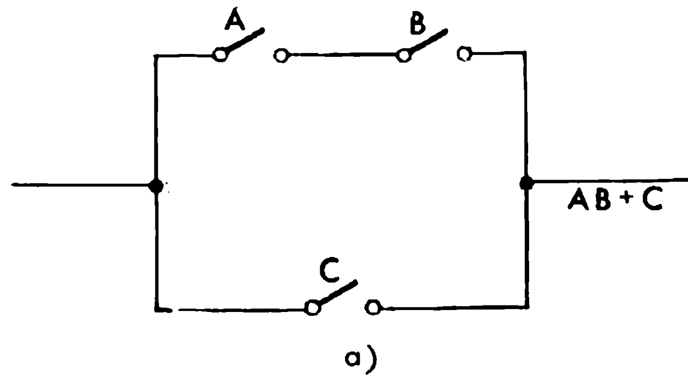


Fig. 31



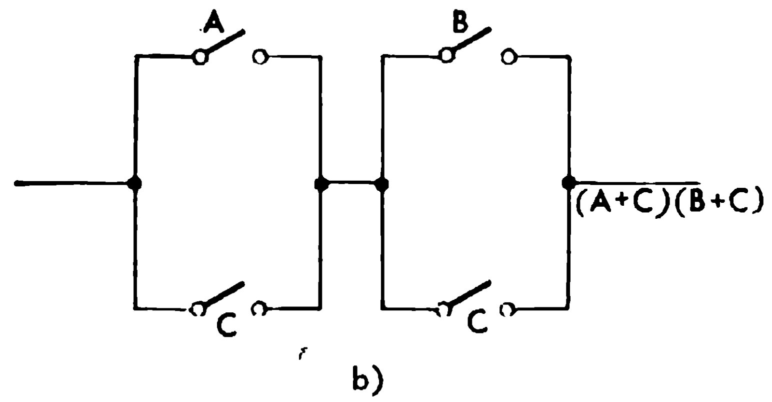
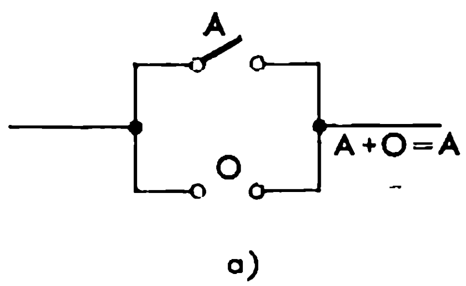


Fig. 32



Fig. 33



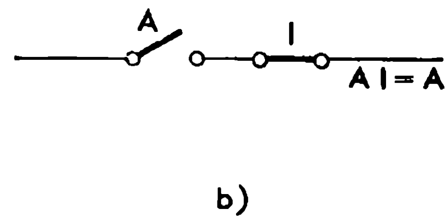


Fig. 34

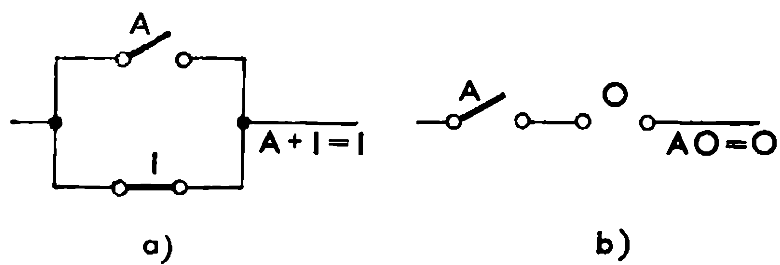


Fig. 35

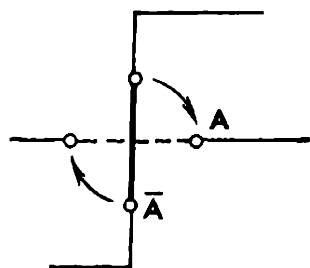
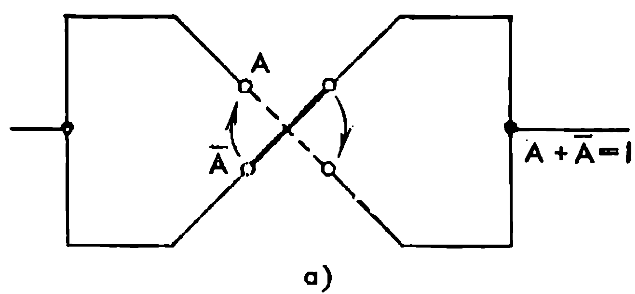


Fig. 36



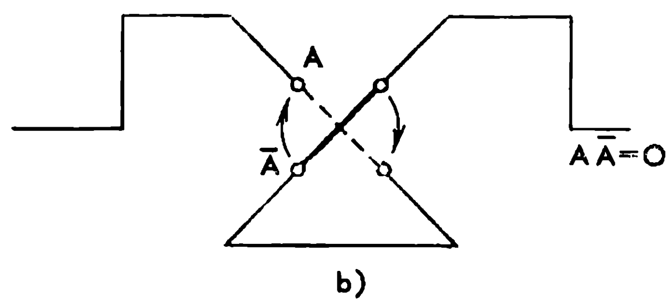
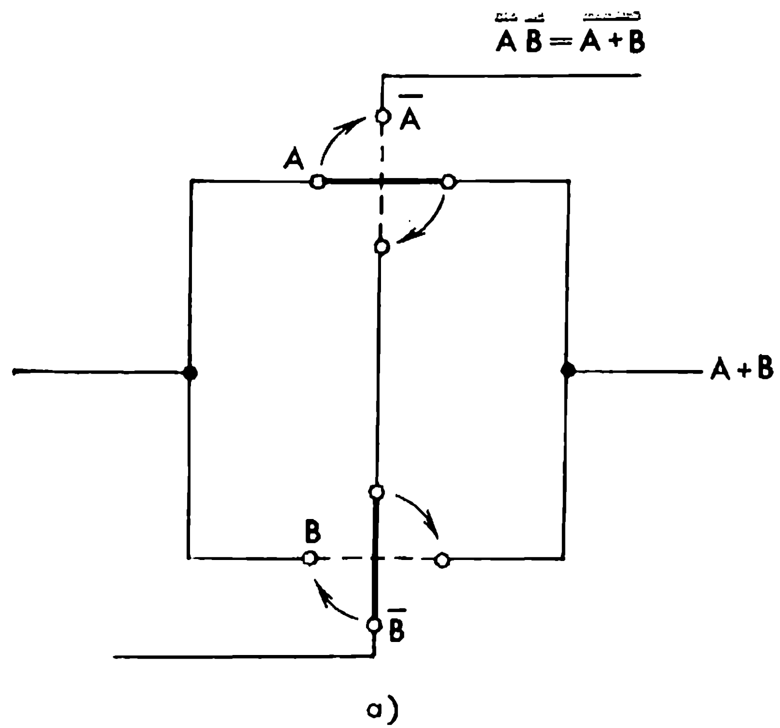


Fig. 37


$\overline{A}\,\overline{B} = \overline{A+B}$
$\overline{A}$
A
A + B
B
$\overline{B}$
a)

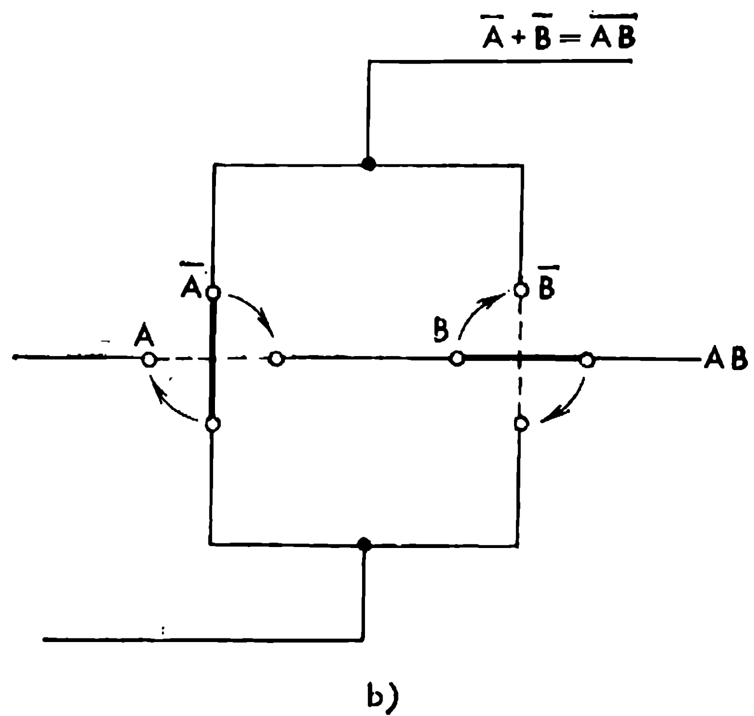

$\overline{A}+\overline{B} = \overline{AB}$
$\overline{A}$
$\overline{B}$
A
B
AB
b)

Fig. 38

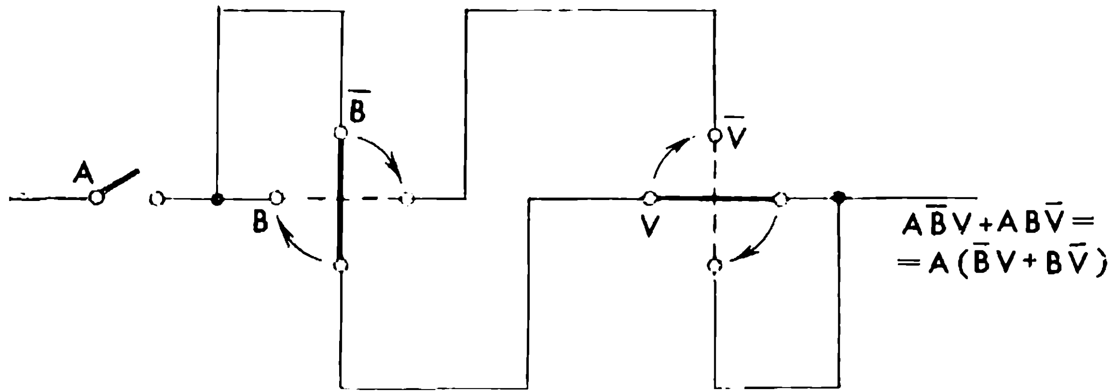


Fig. 39

rapports. Premièrement, elle permet de *simuler* des propositions composées à l'aide de circuits électriques. Examinons, à titre d'exemple, la proposition composée suivante

$$d = a\bar{b}c + ab\bar{c},$$

où a, b et c sont des propositions simples quelconques, tandis que l'addition et la multiplication des propositions, ainsi que l'opération « barre » désignent, comme toujours, les liaisons logiques « ou », « et » et la négation. Faisons correspondre aux propositions a, b et c les contacts A, B et V; dans ce cas la proposition composée d sera représentée par le schéma de la figure 39 qui répond à la combinaison suivante

$$D = A\bar{B}V + AB\bar{V}$$

des contacts A, B, V. Il est à remarquer que pour vérifier si la proposition d est vraie ou non au cas où les propositions a et b sont vraies et la proposition c est fausse, il suffira de fermer les contacts A et B du schéma D et d'ouvrir le contact V (voir fig. 40). Si le schéma D laisse passer le courant, cela signifie qu'il correspond à la proposition vraie i (c'est-à-dire au schéma I qui laisse passer le courant), en d'autres termes, la proposition d est dans ce cas-là vraie. Par contre, si pour les mêmes conditions, le schéma D ne laisse pas passer le

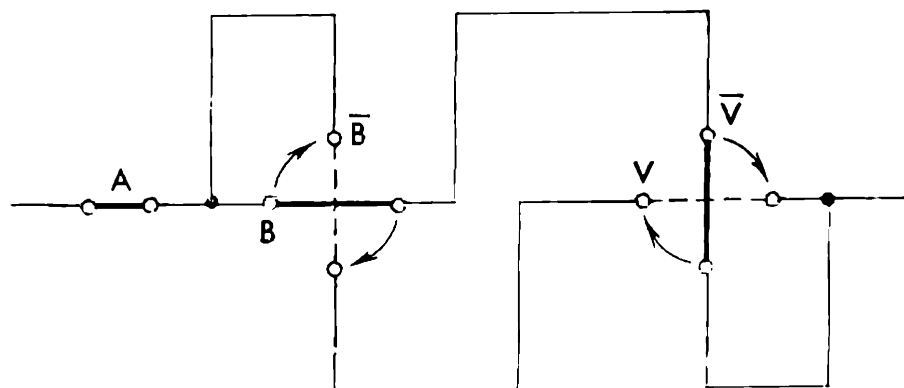


Fig. 40

courant (« équivaut » au schéma O), alors la proposition d équivaut, dans les mêmes conditions, à la proposition o, c'est-à-dire qu'elle est fausse.

Deuxièmement, elle permet, en faisant appel aux règles logiques, de construire des chaînes de contacts (souvent assez compliquées), satisfaisant à des conditions données à l'avance. Et voici deux exemples.

● EXEMPLE 1. On demande de construire un circuit électrique pour une chambre à coucher éclairée par une seule lampe où il est souhaitable d'avoir deux interrupteurs: l'un près de la porte, l'autre près du lit fonctionnant de façon que chaque interrupteur, indépendamment de la position de l'autre, allume ou éteint la lumière.

● SOLUTION. Désignons par A et B les contacts correspondant aux deux interrupteurs; le problème consiste à construire une combinaison V (répondant au circuit électrique de la chambre à coucher) des contacts A et B (et éventuellement $\bar{A}$ et $\bar{B}$) telle que le *changement de la position de l'un quelconque des deux contacts entraine nécessairement la modification de tout le circuit* V (c'est-à-dire ouvre le circuit s'il est fermé et inversement). En d'autres termes, il nous faut trouver une proposition composée c telle qu'en substituant à la proposition vraie a la proposition fausse b et inversement on modifie le caractère de toute la

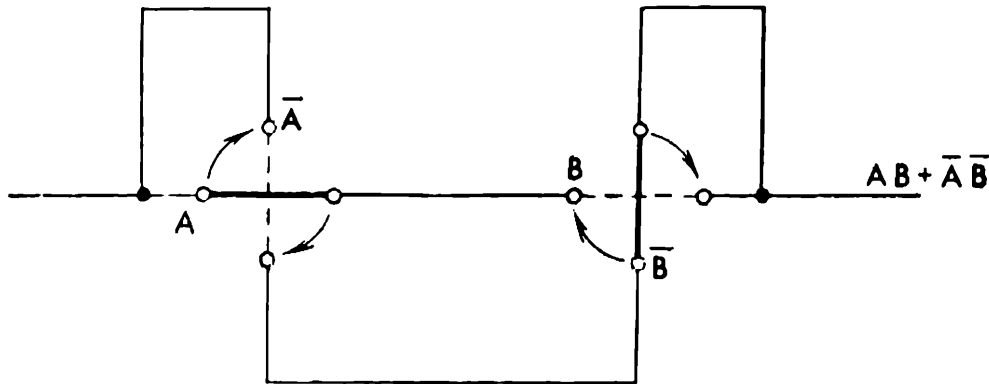


Fig. 41

proposition c (c'est-à-dire de « vraie elle devient fausse » et inversement). Cette condition est vérifiée par la proposition c qui est vraie si les propositions a et b sont *toutes les deux* vraies *ou* fausses, et fausse dans tous les autres cas (lorsque l'une des deux propositions a, b est vraie et l'autre fausse). La conjonction « ou » nous suggère l'idée de mettre la proposition c sous la forme d'une *somme* de deux propositions, l'une étant vraie lorsque a et b sont vraies et l'autre, lorsque $\bar{a}$ et $\bar{b}$ sont vraies (c'est-à-dire lorsque a et b sont fausses). Par ailleurs dans notre raisonnement nous avons utilisé la conjonction « et », donc les deux termes de la somme cherchée sont

$$ab \quad \text{et} \quad \bar{a}\bar{b}.$$

Ainsi, on a définitivement

$$c = ab + \bar{a}\bar{b}.$$

Il est aisé de voir que cette proposition c satisfait à toutes les conditions mentionnées ci-dessus.

En passant maintenant des propositions aux chaînes de contacts, on peut conclure que le circuit électrique V qui nous intéresse s'écrit donc

$$V = AB + \bar{A}\bar{B};$$

il est clair que la réalisation de ce schéma (voir fig. 41) ne pose aucune difficulté sur le plan technique.

● EXEMPLE 2 *). On demande de construire un circuit électrique commandant la marche d'un ascenseur où, pour plus de commodité, le nombre d'étages est supposé égal à deux. Le circuit doit comporter deux contacts commandés par des boutons disposés dans la cabine de l'ascenseur (bouton de descente) et près de la porte d'entrée au rez-de-chaussée (bouton d'appel) ; les contacts supplémentaires commandent les portes de l'ascenseur du rez-de-chaussée et du premier étage, à la porte de la cabine, ainsi qu'au plancher de la cabine sur lequel s'exerce le poids des passagers transportés. Le circuit commandant la descente de l'ascenseur **) ne doit se fermer que si l'ascenseur se trouve au premier étage et si, de plus, sont observées les conditions suivantes :

1) *les deux portes de l'ascenseur et la porte de la cabine sont fermées*; *le passager se trouve dans la cabine et appuie sur le bouton de descente ou*

2) *les deux portes de l'ascenseur sont fermées* (*la porte de la cabine étant fermée ou ouverte*) ; *la cabine est vide*; *un usager appuie sur le bouton d'appel.*

● SOLUTION. Désignons les interrupteurs commandant la fermeture du circuit comme suit : P est un interrupteur ne fonctionnant que si la cabine se trouve au premier étage ; D_r et D_1 ne commencent à fonctionner que si les portes de l'ascenseur sont fermées au rez-de-chaussée et au premier étage ; D_c est identique à celui de la porte de la cabine ; Pl est lié au plancher de la cabine et ne fonctionne que sous le poids du passager ; K_d et K_a sont liés au bouton de descente situé dans la cabine et bouton d'appel se trouvant près de la porte de l'ascenseur au rez-de-chaussée. Selon les données du problème le circuit cherché V_d commandant

---

*) Cet exemple est emprunté au livre de I. Polétaïev, *Signal*, Ed. Radio soviétique, 1958, p. 214.

**) Nous n'examinons ici que le circuit commandant la *descente* de l'ascenseur, celui commandant la montée étant parfaitement analogue (voir exercice 6, p. 89).

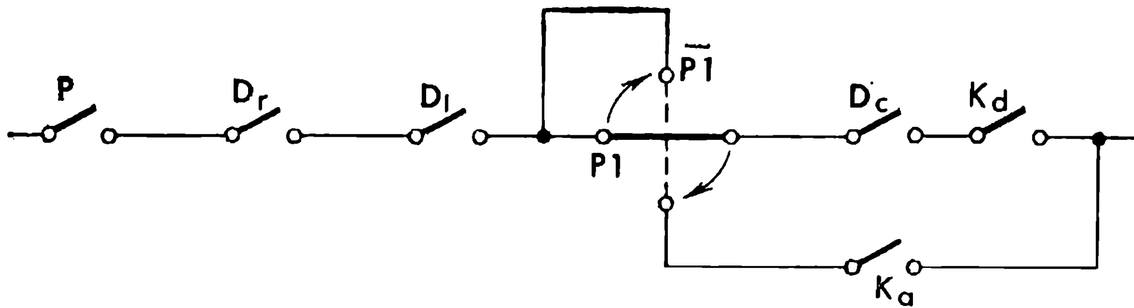


Fig. 42

la descente de l'ascenseur ne doit s'enclencher (c'est-à-dire laisser passer le courant) que si :

1) le contact P est fermé *et* le contact D_r est fermé *et* le contact D_1 est fermé *et* le contact D_c est fermé *et* le contact Pl est fermé *et* le contact K_d est fermé

ou

2) le contact P est fermé *et* le contact D_r est fermé *et* le contact D_1 est fermé *et* le contact D_c est fermé *ou* ouvert, *et* le contact K_a est fermé *et* le contact Pl est ouvert.

Vu que l'opération logique « et » correspond au produit des propositions (ou des contacts) et l'opération logique « ou » à leur somme, on obtient aisément

$$V_d = PD_rD_1D_cPlK_d + PD_rD_1(D_c + D_c)K_aPl.$$

Simplifions cette expression en utilisant l'égalité

$$D_c + D_c = I,$$

la propriété du contact I ($AI = A$ quel que soit A) ainsi que la loi commutative pour la multiplication et la loi distributive. On obtient donc

$$V_d = PD_rD_1(PlD_cK_d + \overline{P}lK_a).$$

La réalisation technique d'un tel circuit ne pose aucune difficulté (voir fig. 42).

Exercices

1. Représenter les chaînes de contacts correspondant aux propositions composées :

a) $(a+b)(c+d)$;

b) $abc+a\bar{b}+\bar{a}$;

c) $ab\bar{c}+a\bar{b}c+\bar{a}bc$;

d) $(a+b)(\bar{a}+\bar{b})+ab+\bar{a}\bar{b}$.

2. Représenter les chaînes de contacts correspondant aux propositions suivantes

$$(a + c)(b + c)(a + d)(b + d) \quad \text{et} \quad ab + cd,$$

vérifier l'« égalité » de ces chaînes.

3*. Construire un circuit électrique V avec les contacts A, B, C et D (et éventuellement les contacts $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{D}$) tel que

a) le circuit V sera fermé seulement si *tous* les contacts A, B, C, D sont fermés ou *aucun* de ces contacts n'est fermé ;

b) le circuit V sera fermé seulement dans le cas où *certains* mais pas tous les contacts A, B, C et D sont fermés.

4. a) Un comité comprend trois membres. On demande de construire un circuit électrique servant à signaler les résultats des votes : en votant, chacun des membres du comité appuie sur un bouton ; la lampe ne s'allume que si la proposition à voter recueille la majorité des voix.

b) Construire un circuit analogue pour un comité qui se compose d'un président et de cinq membres : la lampe ne s'allume que si la proposition à voter recueille la majorité des voix ou la moitié des voix, y compris celle du président.

5*. Construire un circuit électrique permettant d'allumer et d'éteindre une lampe électrique à l'aide de

a) trois commutateurs indépendants (cf. exemple 1, p. 85) ;

b) n commutateurs indépendants.

6. Construire d'après les données de l'exemple 2, p. 87, un circuit commandant la montée d'un ascenseur.

ANNEXE

DÉFINITION DE L'ALGÈBRE DE BOOLE

On appelle *algèbre de Boole* un ensemble arbitraire d'éléments α, β, γ, . . . pour lesquels sont définies deux opérations, l'addition et la multiplication, faisant correspondre à tout couple d'éléments α et β leur *somme* $\alpha + \beta$ et leur *produit* $\alpha\beta$ *); l'opération « barre » qui fait correspondre à chaque élément α un élément nouveau $\overline{\alpha}$ **); il existe deux éléments « particuliers » o et ι et sont satisfaites les règles suivantes:

Règles relatives à l'addition Règles relatives à la multiplication

1) $\alpha+\beta=\beta+\alpha$, 1a) $\alpha\beta=\beta\alpha$;

lois commutatives

2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$, 2a) $(\alpha\beta)\gamma=\alpha(\beta\gamma)$;

lois associatives

3) $\alpha+\alpha=\alpha$, 3a) $\alpha\alpha=\alpha$.

lois idempotentes

Règles qui associent l'addition et la multiplication

4) $(\alpha+\beta)\gamma=\alpha\gamma+\beta\gamma$, 4a) $\alpha\beta+\gamma=(\alpha+\gamma)(\beta+\gamma)$.

lois distributives

Règles relatives aux éléments o *et* ι

5) $\alpha + o = \alpha$, 5a) $\alpha\iota = \alpha$,

6) $\alpha + \iota = \iota$, 6a) $\alpha o = o$.

---

*) Cf. avec la définition de la page 30.

**) Les mathématiciens disent souvent à ce propos qu'en algèbre de Boole il existe deux opérations *binaires* (addition et multiplication) faisant correspondre à chaque couple d'éléments α et β les éléments nouveaux $\alpha + \beta$ et $\alpha\beta$ respectivement et une opération *unaire* qui fait correspondre à chaque élément α un élément nouveau $\overline{\alpha}$.

Règles relatives à l'opération « barre »

7) $\bar{\bar{\alpha}} = \alpha$, 8) $\bar{o} = \iota$, 8a) $\bar{\iota} = o$.

Règles liant l'opération « barre » à l'addition et à la multiplication

9) $\overline{\alpha + \beta} = \bar{\alpha}\bar{\beta}$, 9a) $\overline{\alpha\beta} = \bar{\alpha} + \bar{\beta}$.

Règles de De Morgan

Dans la définition de l'algèbre de Boole il n'est pas nécessaire d'exiger qu'ait lieu la relation $\supset$: l'inclusion $\alpha \supset \beta$ peut bien *être définie* et ses propriétés déduites à partir de n'importe laquelle des conditions $\alpha + \beta = \alpha$ et $\alpha\beta = \beta$

$$\alpha \supset \alpha;$$
$$\text{si } \alpha \supset \beta \text{ et } \beta \supset \alpha, \quad \text{alors} \quad \alpha = \beta;$$
$$\text{si } \alpha \supset \beta \text{ et } \beta \supset \gamma, \quad \text{alors} \quad \alpha \supset \gamma;$$
$$\iota \supset \alpha \quad \text{et} \quad \alpha \supset o;$$
$$\alpha + \beta \supset \alpha \quad \text{et} \quad \alpha \supset \alpha\beta;$$
$$\text{si } \alpha \supset \beta, \quad \text{alors} \quad \bar{\beta} \supset \bar{\alpha}$$

(nous laissons au lecteur le soin de les déduire). Bien plus, dans la définition de l'algèbre de Boole il n'est pas obligatoire d'exiger l'existence d'une des deux opérations: d'addition ou de multiplication, mais notamment de la deuxième de ces opérations et de l'opération « barre »; ainsi, les opérations « addition » et « barre » permettent de définir la multiplication grâce aux règles de De Morgan:

$$\alpha\beta = \overline{\bar{\alpha} + \bar{\beta}}.$$

Cependant les opérations d'addition et de multiplication (l'opération « barre » non comprise) ne constituent pas encore par elles-mêmes l'algèbre de Boole.

La définition de l'algèbre de Boole énoncée ci-dessus est loin d'être « rationnelle » en ce sens que beaucoup d'opérations mentionnées peuvent être déduites des autres.

RÉPONSES
ET INDICATIONS

§ 1

1. $(A+B)(A+C)(B+D)(C+D) = [(B+A)(C+A)]\times$ $\times[(B+D)(C+D)] = (BC+A)(BC+D) = (A+BC)(D+BC) =$ $= AD+BC$ (on se sert ici de la deuxième loi distributive).

2. $A(A+B) = AA+AB = A+AB = AI+AB = A(I+B) =$ $= AI = A$.

5. $A(A+I)(B+O) = A\cdot I\cdot B = AB$.

6. $(A+B)(B+C)(C+A) = ABC+AB+AC+BC = ABC+$ $+ABI+AC+BC = AB(C+I)+AC+BC = ABI+AC+$ $+BC = AB+BC+CA$
(voir l'identité démontrée à la page 28).

7. $[(A+B)(B+C)](C+D) = (AC+B)(C+D) = AC+ACD+$ $+BC+BD = AC+BC+BD$.

10. $[(A+B+C)(B+C+D)](C+D+A) = [AD+(B+C)]\times$ $\times(C+D+A) = [(AD+B)+C][(A+D)+C] = (AD+B)\times$ $\times(A+D)+C = AD+AD+AB+BD+C = AB+AD+$ $+BD+C$.

§ 2

3. a)

$+$	O	I
O	O	I
I	I	I

et

$\cdot$	O	I
O	O	O
I	O	I

b)

+	O	P	Q	I
O	O	P	Q	I
P	P	P	I	I
Q	Q	I	Q	I
I	I	I	I	I

et

·	O	P	Q	I
O	O	O	O	O
P	O	P	O	P
Q	O	O	Q	Q
I	O	P	Q	I

6. $[m, n] = p_1^{\max[a_1, b_1]} p_2^{\max[a_2, b_2]} \ldots p_k^{\max[a_k, b_k]}$;

$(m, n) = p_1^{\min[a_1, b_1]} p_2^{\min[a_2, b_2]} \ldots p_k^{\min[a_k, b_k]}$.

§ 3

1. $AB + AC + BD + CD = (A + D)(B + C)$ (voir exercice 1 § 1); $A + AB = A$ (voir exercice 2 § 1); $AB + BO + AI = A$ (voir exercice 9 § 1); $ABC + BCD + CDA = (A + B)(A + D)(B+D)C$ (voir exercice 10 § 1).

2. a) $(A+B)(A+\bar{B}) = AA + A\bar{B} + BA + B\bar{B} = A + A\bar{B} + BA + O = A + BA + \bar{B}A = A + (B + \bar{B})A = A + IA = A + A = A$;

b) $AB + (A+B)(\bar{A}+\bar{B}) = AB + A\bar{A} + A\bar{B} + \bar{B}A + B\bar{B} = AB + O + A\bar{B} + B\bar{A} + O = AB + A\bar{B} + B\bar{A} = (AB + A\bar{B}) + (AB + \bar{A}B) = A(B + \bar{B}) + (A + \bar{A})B = AI + IB = A + B$;

c) $\overline{\bar{A}BC}\,\overline{\bar{A}B}\,\bar{A}C = (\bar{\bar{A}} + \bar{\bar{B}} + \bar{C})(\bar{\bar{A}} + \bar{B})\bar{A}C = [(A+B)+\bar{C}](A+\bar{B})\bar{A}C = A(A + \bar{B})\bar{A}C + B(A + \bar{B})\bar{A}C + (A+\bar{B})\bar{A}(C\bar{C}) = (A\bar{A})(A+\bar{B})C + [B(A\bar{A})C + (B\bar{B})\bar{A}C] + (A+\bar{B})\bar{A}O = O(A+\bar{B})C + [BOC + O\bar{A}C] + O = O$;

d) $A + B = A + IB = A + (A + \bar{A})B = A + AB + \bar{A}B = (AI + AB) + \bar{A}B = A(I + B) + \bar{A}B = AI + \bar{A}B = A + \bar{A}B$.

3. Appliquer aux deux membres de l'égalité considérée l'opération « barre »; se servir de l'égalité $\bar{\bar{A}} = A$.

4. $AB + A\bar{B} = A$ (voir exercice 2a); $(A + B)(AB + \bar{A}\bar{B}) = AB$ (voir exercice 2b); $A(\bar{A} + B) = AB$ (voir exercice 2d).

6. a) A chaque diviseur m du nombre N correspond un certain sous-ensemble de l'ensemble $I = \{p_1, p_2, \ldots, p_k\}$ des diviseurs premiers du nombre N, c'est-à-dire l'ensemble de ceux de ces diviseurs qui le sont en même temps pour le nombre m; si de plus aux nombres m et n correspondent les sous-ensembles A et B de l'ensemble I, alors aux nombres $m \oplus n = [m, n]$, $m \otimes n = (m, n)$ et $\bar{m} = \frac{N}{m}$ correspondent les ensembles $A + B$, AB et $\bar{A}$. b) Si $m = p^a$, $n = p^b$, alors $m \oplus n = [m, n] = p^{\max[a, b]}$, $m \otimes n = (m, n) = p^{\min[a, b]}$ et $\bar{m} = \frac{N}{m} = p^{A-a}$. c) $\bar{m} = \frac{N}{m} = p_1^{A_1-a_1} p_2^{A_2-a_2} \ldots p_k^{A_k-a_k}$.

7. Dans l'« algèbre des maxima et des minima » ces égalités ne sont pas satisfaites (sauf si cette algèbre n'est composée que de deux nombres quelconques tout au plus); il en est de même pour l'« algèbre des plus petits multiples et des plus grands diviseurs » (sauf le cas où tous les diviseurs premiers $p_1, p_2, \ldots, p_k$ du nombre N sont distincts deux à deux; cf. exercice 6a).

8. a) $(A + B)(A + C) = A + AC + AB + BC = AI + AC + AB + BC = A(I + C + B) + BC = AI + BC = A + BC \subset A + B \subset A + B + C$;

b) $(A + B)(A + C)(A + I) = (A + B)(A + C) I = (A+B) \times (A + C) = A + BC \supset A \supset ABC$ (cf. exercice a);

c) $(A + B)(B + C)(C + A) = AB + BC + CA \supset AB \supset ABC$ (voir exercice 6 § 1);

d) vu que $A \supset A\bar{B}$ et $B \supset \bar{A}B$, alors $A + B \supset A\bar{B} + \bar{A}B$.

9. $ABC \subset AB + AC$ (voir exercice 8a); $AB + AC + AO \subset A + B + C$ (voir exercice 8b); $AB + BC + CA \subset A + B + C$ (voir exercice 8c).

10. $AB \subset (\bar{A} + B)(A + \bar{B})$.

12. a) A; b) C; c) I; d) O.

§ 4

5. a) « il est pair *et* il est premier »; l'ensemble de vérité est $\{2\}$; b) « il est impair *ou* il est premier »; l'ensemble de vérité $\{1, 2, 3, 5, 7, 9, 11, 13, 15, 17 \ldots\}$ diffère de l'ensemble des nombres impairs

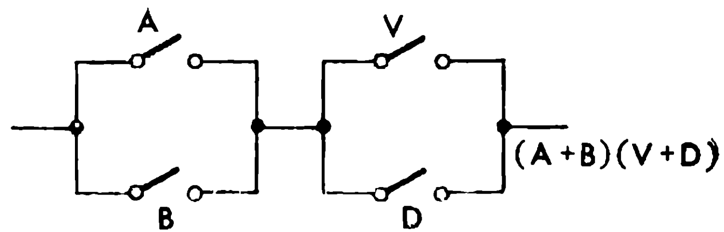


Fig. 43

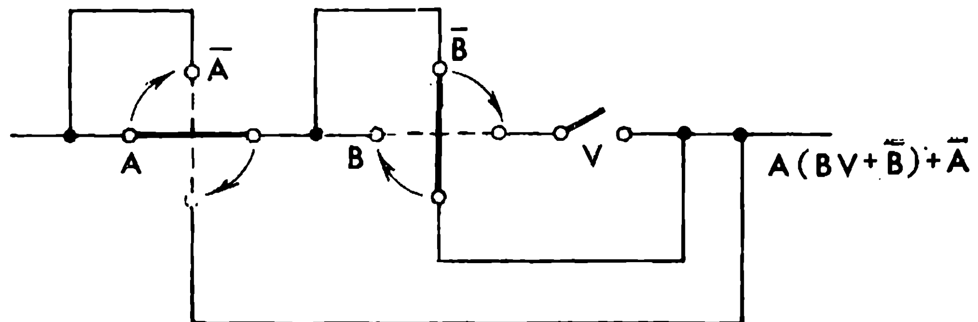
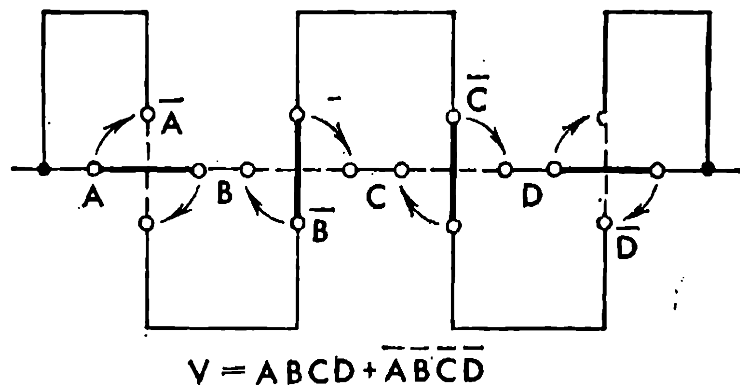


Fig. 44



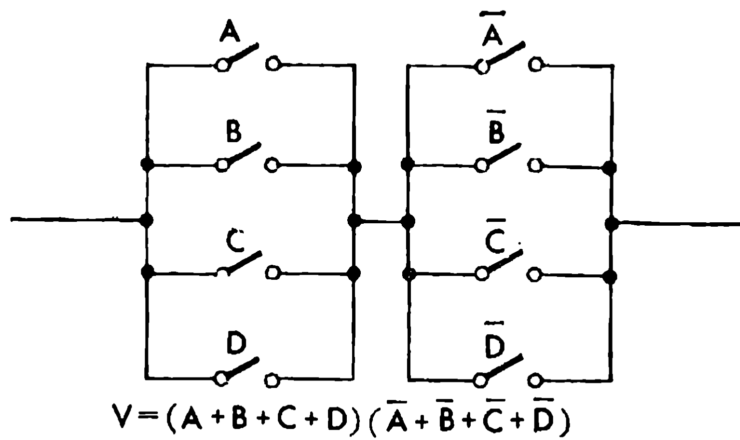


Fig. 45

par l'adjonction du nombre 2; c) « il est impair *et* il est premier »; l'ensemble de vérité {3, 5, 7, 11, 13, 17, 19 . . .} diffère de l'ensemble des nombres premiers par l'élimination du nombre 2; d) « il est pair *et* il n'est pas premier »; l'ensemble de vérité {4, 6, 8, 10, 12, 14, 16...} se distingue de l'ensemble des nombres pairs par l'élimination du nombre 2; e) « il est impair *ou* il n'est pas premier »; l'ensemble de vérité {1, 3, 4, 5, 6, 7, 8, 9, 10, 11 . . .} est l'ensemble de tous les nombres entiers positifs, sauf le nombre 2.

§ 5

5. $a + b = a$; $ab = b$.

§ 4

1. a) Voir fig. 43; b) voir fig. 44. 3. a) Voir fig. 45,*a*; b) voir fig. 45,*b*.
4. a) $C = AB + AC + BC$;
b) $C = A\,(BC + BD + BE + BF + CD +$
$+ CE + CF + DE + DF + EF) + BCDE + BCDF + BCEF +$
$+ BDEF + CDEF$ (le bouton A est appuyé par le président du comité).

5. a) $V = ABC + A\overline{B}\overline{C} + \overline{A}B\overline{C} + A\overline{B}\overline{C}$.

B. Trakhtenbrot

ALGORITHMES ET RÉSOLUTION DE PROBLÈMES PAR DES MACHINES

INTRODUCTION

Dans les années d'après-guerre se sont beaucoup développées les machines à calculer rapides qui sont à présent utilisées pour résoudre les problèmes mathématiques et logiques les plus divers. La singularité caractéristique de ces machines, qui les distingue des calculatrices plus anciennes, consiste en ce que, lorsqu'elles remplissent leurs fonctions, à partir du moment où elles ont reçu les données initiales et le programme, elles fonctionnent jusqu'à la délivrance du résultat final sans que l'homme ait à intervenir. La productivité des machines électroniques et automatiques modernes est énorme : elles peuvent faire jusqu'à 20 000 opérations arithmétiques par seconde, soit au moins 10 fois plus qu'un calculateur qualifié opérant sur un bon arithmomètre à clavier au cours d'une journée de travail *). Le domaine d'utilisation des machines automatiques continue à s'étendre : celles-ci résolvent des systèmes complexes d'équations, traduisent un langage dans un autre, jouent aux échecs, etc. Les machines automatiques ont de belles perspectives d'application à la production, en assurant le contrôle de tout le processus technologique à l'échelle d'une grande usine. En outre, la possibilité d'une mise au point mathématique rapide et sûre, ainsi que l'analyse des données expérimentales constituent les prémisses à l'apparition dans de nombreux domaines de la science de nouvelles méthodes de recherche jusque-là inaccessibles.

Nul doute aujourd'hui que les machines à calculer automatiques représentent un instrument puissant du travail intellectuel, capable non seulement de faciliter le travail intellectuel humain, mais encore de libérer complètement l'homme de certains aspects pesants et contraignants de ce travail.

---

*) Du point de vue de l'exécution de ces opérations.

D'autre part, les succès obtenus peuvent faire naître — et c'est le cas — beaucoup d'illusions non justifiées et de pronostics purement fantasques à propos de la toute-puissance des machines. Il convient surtout de mentionner le battage publicitaire né dans une partie de la presse étrangère au sujet du « gigantesque cerveau électronique » et des automates capables de résoudre n'importe quel problème et de remplacer le travail créateur du savant.

De ce point de vue, la question de savoir de quels aspects du travail intellectuel peuvent se charger les machines à calculer automatiques prend une grande actualité et se pose avec beaucoup d'acuité. Cette question est examinée et résolue d'une certaine façon dans la théorie moderne des algorithmes, qui constitue une branche importante de la logique mathématique. Celle-ci se caractérise par la recherche de la nature de notions telles que « processus calculatoire », « démonstration mathématique », « algorithme », etc. Quelques années encore avant l'apparition des machines à calculer automatiques et électroniques modernes, on avait mis au point, en logique mathématique, la notion exacte d'« algorithme » ainsi qu'un schéma général de machine à calculer automatique (machine de Turing); on avait d'autre part mis en évidence le lien étroit entre les algorithmes et les machines. Cela a permis d'établir une série de théorèmes importants qui dévoilent la nature des processus réalisés dans les machines automatiques; en particulier, on a démontré de façon rigoureuse l'existence de problèmes impossibles à résoudre par les machines. Le présent ouvrage est consacré à l'étude du lien entre les algorithmes et les machines.

Dans les §§ 1-3, on explique sur une série d'exemples ce que signifie un algorithme et on construit des algorithmes pour résoudre certains types de problèmes mathématiques et logiques.

Dans les §§ 4-5 sont exposés les principes de construction des machines à calculer électroniques et d'élaboration des

programmes, i.e. des algorithmes capables de réaliser ces derniers dans les machines.

Les §§ 6-11 sont consacrés à une série d'éléments importants de la théorie des algorithmes, la machine de Turing étant prise comme notion fondamentale de cette théorie.

La longueur purement technique de beaucoup de démonstrations ne permet pas d'en faire l'exposé détaillé dans un si petit ouvrage. C'est pourquoi on s'écarte un peu de la rigueur et l'exposé n'est pas complet ; mais il nous a semblé que cela n'entravait pas, mais au contraire favorisait la mise en lumière de la nature du problème. Pour brosser un tableau général, un tour d'horizon est fait au § 6.

Nous ferons encore une remarque. Les machines modernes à commande automatique sont appelées électroniques du fait que leurs organes essentiels sont constitués de lampes électroniques. L'application de la technique électronique permet une économie substantielle de temps lors de la réalisation des opérations isolées effectuées par la machine. Mais la particularité fondamentale de ces machines, à savoir la commande automatique des processus dont elles sont le siège, ne ressort pas particulièrement de la technique électronique. Les lampes électroniques auraient pu en principe être remplacées par des organes même mécaniques, i.e. on aurait pu construire une machine à calculer mécanique à commande automatique capable de résoudre (bien que beaucoup plus lentement) les mêmes problèmes que la machine électronique. Ainsi, l'apparition de machines à calculer d'un type nouveau ne peut pas être considérée comme le résultat du développement de la seule technique électronique. De plus, la première description d'un schéma général de machine à calculer automatique était encore donnée en 1936, dans la théorie des algorithmes, comme celle d'un assemblage mécanique (machine de Turing, cf. n° 7). En fait, les premières machines à calculer (1940) étaient électromécaniques.

Lors de la description de la structure des machines, nous ne nous arrêterons pas dans cet ouvrage aux détails techniques, mais accorderons une attention particulière à l'étude des principes d'interaction des organes de la machine. Une telle démarche correspond au but essentiel du livre qui est de montrer les possibilités mathématiques et logiques des machines et non l'aspect technique du problème.

§ 1. ALGORITHMES NUMÉRIQUES

La notion d'algorithme figure parmi les notions fondamentales des mathématiques. Par *algorithme* on entend un programme d'opérations que la machine doit effectuer dans un certain ordre et qui permettent de résoudre tous les problèmes d'un type déterminé.

Naturellement, cette phrase n'est pas la définition mathématique rigoureuse d'*algorithme*; elle explique plutôt le sens du mot « *algorithme* » et met en lumière sa signification. Une telle formulation n'en est pas moins familière et accessible à tout mathématicien; elle reprend la notion d'algorithme qui a pris spontanément naissance et est utilisée en mathématiques depuis les temps les plus reculés.

Les algorithmes les plus simples sont des règles suivant lesquelles sont effectuées l'une ou l'autre des quatre opérations arithmétiques dans le système décimal (le terme lui-même d'*algorithme* vient du nom du mathématicien ouzbek du Moyen-Age Al-Khorezmi, qui énonça ces règles dès le IXe siècle). Ainsi, par exemple, l'opération d'addition de deux nombres comportant plusieurs chiffres se décompose en une chaîne d'opérations élémentaires; lorsqu'il effectue chacune d'entre elles, le calculateur ne considère que les deux chiffres correspondants des nombres à ajouter (l'un d'entre eux peut comporter une retenue). Ces opérations sont de deux types: 1) inscription du chiffre correspondant de la somme, 2) retenue quand on passe au chiffre situé immédiate-

ment à gauche du précédent; en plus, cette règle indique l'ordre dans lequel doivent être effectuées les opérations (de droite à gauche). Le caractère formel de ces opérations élémentaires consiste en ce qu'elles peuvent être effectuées automatiquement d'après une table d'addition des chiffres donnée une fois pour toutes, en faisant complètement abstraction de leur signification.

Il en va de même pour les trois autres opérations arithmétiques, ainsi que pour l'opération d'extraction de racine carrée, etc. Le caractère formel des programmes correspondants (algorithmes) ne fait apparemment aucun doute (pour les écoliers, cela est surtout sensible dans la règle d'extraction des racines).

A titre d'exemple, considérons l'*algorithme d'Euclide*, qui résout tous les problèmes du type suivant:

Etant donné deux nombres naturels a et b, trouver leur plus grand commun diviseur.

Il existe évidemment autant de problèmes distincts de ce type qu'il existe de couples différents de nombres (a, b).

On sait qu'il est possible de résoudre n'importe lequel de ces problèmes en construisant une suite décroissante de nombres, le premier étant le plus grand des deux nombres donnés, le second le plus petit; le troisième est obtenu comme le reste de la division du premier par le second, le quatrième, comme le reste de la division du second par le troisième, etc., jusqu'à ce qu'on obtienne une division dont le reste est nul. Le diviseur, dans cette dernière division, sera le résultat cherché.

Dans la mesure où la division se ramène à la répétition de soustractions, on pourrait donner un programme adapté à la résolution de n'importe lequel de ces problèmes sous la forme de la série d'instructions suivante:

Instruction 1. Examiner deux nombres a et b. Passer à l'instruction suivante.

Instruction 2. Comparer les nombres considérés ($a = b$, ou $a < b$, ou $a > b$); passer à l'instruction suivante.

Instruction 3. Si les nombres en question sont égaux, alors chacun d'entre eux fournit le résultat cherché. Arrêter le processus de calcul. Sinon, passer à l'instruction suivante.

Instruction 4. Si le premier nombre est inférieur au second, les permuter et passer à l'instruction suivante.

Instruction 5. Retrancher le second des deux nombres obtenus du premier et considérer les deux nombres: le nombre soustrait et le reste. Passer à l'instruction 2.

Ainsi, après avoir accompli ces 5 instructions, il faut revenir à la seconde, puis à la troisième, quatrième et cinquième, puis à nouveau à la seconde, troisième, etc., jusqu'à ce qu'on obtienne des nombres égaux, i.e. jusqu'à ce que la condition stipulée dans l'instruction 3 soit remplie; alors le processus s'arrête.

Il est vrai qu'en mathématiques les algorithmes ne se formulent pas toujours sous une forme aussi pédante; mais il ne fait aucun doute qu'on puisse toujours formuler de cette façon n'importe lequel des algorithmes connus.

Dans la description précédente de l'algorithme d'Euclide, les opérations élémentaires dont se compose le processus de résolution du problème sont: la soustraction de deux nombres, leur comparaison et leur permutation; mais il est aisé de comprendre qu'on puisse pousser la décomposition beaucoup plus loin. Ainsi, par exemple, l'instruction 5 qui consiste à soustraire les deux nombres considérés peut se décomposer elle-même en un système d'instructions décrivant l'algorithme de soustraction des deux nombres. Cependant, comme les règles des opérations arithmétiques dans de tels cas sont très simples et habituelles, on ne détaillera pas davantage l'algorithme.

Les algorithmes qui correspondent à des problèmes dont la résolution se ramène aux quatre opérations arithmétiques sont appélés *algorithmes numériques.* Ils jouent un rôle important dans les domaines les plus divers des mathématiques, tant élémentaires que supérieures, et s'énoncent souvent sous la forme d'instructions verbales ou de formules et sché-

mas divers. Ainsi, par exemple, l'algorithme de résolution d'un système de deux équations du premier dégré à deux inconnues :

$$a_1x + b_1y = c_1,$$
$$a_2x + b_2y = c_2,$$

est donné par les formules

$$x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1} \; ; \quad y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1},$$

qui expriment parfaitement quelles sont les opérations à effectuer et dans quel ordre elles doivent être faites. Dans les formules précédentes, on prévoit une seule et même chaîne d'opérations pour tous les problèmes d'un type donné (i.e. pour n'importe quels coefficients a_1, b_1, c_1, a_2, b_2, c_2). Il est cependant intéressant de noter que, en général, le nombre des opérations indiquées par l'algorithme n'est pas connu à l'avance ; il dépend du choix concret des conditions du problème et on ne le découvre que dans le processus lui-même de résolution.

Il en va ainsi, en particulier, dans le cas de l'algorithme d'Euclide, où le nombre de soustractions à effectuer dépend du choix du couple (a, b) de nombres.

L'important développement des algorithmes numériques est dû à ce que l'on peut ramener aux quatre opérations arithmétiques un bon nombre d'autres opérations. Bien sûr, une telle réduction n'est pas parfaitement précise, mais on peut l'effectuer avec n'importe quelle précision donnée à l'avance. On pourrait déjà illustrer tout cela sur l'exemple de l'algorithme permettant d'extraire une racine carrée, et qui fournit la racine de manière approchée, mais avec n'importe quelle précision donnée à l'avance, à l'aide d'une série de divisions, de multiplications et de soustractions. Dans une branche particulière des mathématiques modernes (*analyse numérique*), on met au point des méthodes analogues pour

ramener des opérations plus complexes, telles que l'intégration, la différentiation, etc. à des opérations arithmétiques.

En mathématiques, une série de *problèmes d'un type donné est considérée comme résolue si on a établi un algorithme pour sa résolution.* La recherche de tels algorithmes est un objectif naturel des mathématiques. C'est ainsi qu'en algèbre on a mis au point des algorithmes permettant de définir de manière parfaitement automatique le nombre des racines distinctes d'une équation algébrique donnée (et leur ordre de multiplicité), lorsque les coefficients de cette équation sont donnés, et de calculer ces racines avec une précision donnée à l'avance.

Si le mathématicien ne dispose pas d'algorithme pour résoudre tous les problèmes d'un type donné, alors il devra, même s'il arrive parfois à résoudre certains problèmes de ce type, inventer une procédure spéciale dans chaque cas isolé, qui ne conviendra plus alors pour la majorité des autres cas.

Donnons un exemple d'une telle classe de problèmes de même type pour laquelle les mathématiques ne disposent pas actuellement d'algorithme.

Considérons toutes les équations diophantiennes imaginables, c'est-à-dire les équations du type

$$P = 0,$$

où P est un polynôme à coefficients entiers. Telles seront, par exemple, les équations

$$x^2 + y^2 - z^2 = 0,$$
$$6x^{18} - x + 3 = 0,$$

où la première est une équation à trois inconnues, la seconde, à une inconnue (on considère généralement des équations avec un nombre quelconque d'inconnues). L'équation peut avoir ou non une solution entière. Ainsi, la première équation en possède une:

$$x = 3, \ y = 4, \ z = 5;$$

la seconde n'en a pas, car pour tout entier x on établit facilement l'inégalité

$$6x^{18} > x - 3.$$

En 1901, au Congrès International de Mathématiques de Paris, le grand mathématicien allemand David Hilbert donna la liste de 20 problèmes difficiles et attira l'attention des participants sur l'importance de leur résolution. Parmi eux figurait le problème suivant (10-ème problème de Hilbert) : *mettre au point un algorithme qui permettrait de déterminer si une équation diophantienne a ou non une solution entière.*

On connaît depuis longtemps un tel algorithme pour le cas particulier d'une équation diophantienne à une inconnue.

En effet, on a montré que si l'équation à coefficients entiers

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0$$

admet une racine entière x_0, alors a_0 est nécessairement divisible par x_0. On peut proposer donc l'algorithme suivant :

1) trouver tous les diviseurs de a_0 (il y en a un nombre fini) ;

2) remplacer successivement tous les diviseurs trouvés dans le premier membre de l'équation et calculer sa valeur numérique ;

3) si par substitution d'un diviseur le premier membre de l'équation s'annule, alors ce diviseur est la racine de l'équation ; si, par contre, le premier membre ne s'annule pour aucun diviseur, alors l'équation n'a pas de racine entière.

Le problème de Hilbert a retenu jusqu'à maintenant l'attention de grands mathématiciens ; mais même à l'heure actuelle, on ne connaît pas l'algorithme dans le cas général où l'équation est à deux inconnues ou plus. En outre, on a maintenant de bonnes raisons de penser qu'on ne le trouvera jamais. Cependant, le sens exact de ce pronostic à pre-

mière vue pessimiste sera seulement éclairci tout à la fin de cet exposé.

Dans les exemples examinés jusqu'à présent, apparaissent déjà avec suffisamment de netteté les caractéristiques suivantes des algorithmes numériques, qui sont propres à tout autre algorithme, à savoir :

Caractère déterminant d'un algorithme. On impose que la méthode de calcul puisse être communiquée à n'importe qui sous la forme d'un nombre fini d'instructions portant sur la manière de procéder à chacune des étapes de calcul prise isolément. De plus, les calculs effectués conformément à ces instructions ne dépendent pas du calculateur et représentent un processus déterminé qui peut être itéré à n'importe quel moment, et avec le même succès par n'importe qui.

Caractère universel d'un algorithme. Un algorithme est un programme unique qui détermine un processus calculatoire ; celui-ci peut partir de données initiales différentes ; il conduit dans tous les cas au résultat correspondant. En d'autres termes, un algorithme ne résout pas un problème individuel, mais toute une série de problèmes d'un même type.

§ 2. ALGORITHMES POUR LA RÉSOLUTION DE PROBLÈMES LOGIQUES

Les problèmes considérés au paragraphe précédent étaient du domaine de l'arithmétique, de l'algèbre et de la théorie des nombres ; ils sont suffisamment typiques pour ces branches des mathématiques et portent, pour ainsi dire, un caractère mathématique traditionnel. Nous allons maintenant nous occuper de deux classes de problèmes qui offrent un caractère quelque peu différent ; on pourrait les appeler logiques, plutôt que mathématiques, bien qu'il soit très difficile de proposer un critère suffisamment précis pour distinguer les problèmes *logiques* de ce genre des problèmes *mathématiques* classiques. Sans nous lancer par la suite dans un

examen de la terminologie, nous remarquerons simplement qu'il s'agit tout de même dans les deux cas de trouver un algorithme qui permette, par une méthode unique, de résoudre n'importe lequel des problèmes de même type de la classe envisagée; cependant, dans les cas que nous allons examiner, ces algorithmes ne sont plus numériques.

1. *Jeu du « 15 »*. Un tableau carré est divisé en seize cases égales. Dans quinze d'entre elles, choisies arbitrairement, on peut disposer quinze fiches numérotées de 1 à 15 (une par case) et de manière arbitraire; on forme ainsi une *position*. Nous dirons que deux cases sont *voisines* si elles ont un segment commun. Deux positions seront considérées comme *contiguës* si l'on peut passer de l'une à l'autre par une seule *marche* qui consiste à placer dans la case vide l'une quelconque des fiches provenant d'une case voisine.

Le nombre de toutes les positions possibles est fini et vaut

$$16! = 20\,922\,789\,888\,000;$$

de plus, à partir d'une position donnée, on ne peut effectuer qu'un nombre fini de marches (2, 3 ou 4).

Il se pose la série suivante de problèmes du même type:

Etant donné deux positions A et B, est-il possible de passer de l'une à l'autre par un nombre fini de marches?

Si la réponse à la question posée est affirmative, alors la suite des marches correspondantes donne naissance à une chaîne de positions:

$$A \leftrightarrow A_1 \leftrightarrow A_2 \leftrightarrow \ldots \leftrightarrow A_n \leftrightarrow B,$$

qui va de A à B, où $\leftrightarrow$ désigne une marche transformant la position donnée en une position contiguë. On peut considérer qu'on ne retrouve pas deux fois la même position dans cette chaîne, car sinon on pourrait supprimer toute la portion de la chaîne qui relie ces deux positions identiques et obtenir par là même une chaîne plus courte allant de A à B.

Ainsi, si la réponse est affirmative, le nombre de marches nécessaires ne dépasse pas 16! — 1. A partir de là, on peut maintenant formuler pour le problème posé l'algorithme suivant qui repose simplement sur l'idée de traitement de toutes les combinaisons possibles de 1, 2, 3, . . ., 16!—1 marches. On dresse la liste des positions, dans laquelle on inclut: la position A, les positions contiguës par rapport à A, les positions contiguës par rapport à ces dernières, etc., 16!—1 fois. Si dans cette liste on rencontre B, alors la réponse est affirmative, sinon elle est négative.

La méthode de résolution proposée est très lourde. Mais pour ce qui est de savoir si l'on a un programme précis qui permette de résoudre, par une méthode unique et à la suite d'un nombre fini d'étapes, tout problème du même type, nous devons répondre par l'affirmative. Nous entendons évidemment par là que le processus décrit est possible à réaliser (sa réalisation pratique dépend des moyens mis à la disposition du calculateur); il est fini, bien que très long.

Quoique cet algorithme soit peu attrayant, le seul fait de l'avoir trouvé est important. Car pour le problème de Hilbert sur les équations diophantiennes, on n'a pas réussi à trouver le moindre algorithme! Par ailleurs, la mise en place d'un algorithme, aussi lourd soit-il, peut laisser espérer en son amélioration ou en la construction d'algorithmes plus maniables. C'est en particulier ce qui se passe dans le cas que nous envisageons. Nous allons maintenant décrire un autre algorithme, qui est tout à fait acceptable du point de vue de sa maniabilité; mais nous aurons besoin pour le justifier d'un théorème que nous énoncerons plus bas sans démonstration.

Nous appellerons *transposition* un échange de deux fiches sur le tableau. Remarquons que l'on peut passer d'une position A à une position B au moyen de déplacements et de transpositions; si les cases vides ne coïncident pas dans ces deux positions, on transforme d'abord B en une position B'

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

Position A

10	1	4	5
9	2	6	8
11	3		15
13	14	7	12

Position B

10	1	4	5
9	2	6	8
11	3	15	12
13	14	7	

Position C

Fig. 1

en effectuant seulement des déplacements (pas plus de 15), et où B' présente la même case vide que A. Ensuite, le passage de B' à A s'effectue à l'aide des seules transpositions (pas plus de 15 transpositions). L'exactitude de notre affirmation devient évidente si l'on considère un exemple. Supposons qu'il s'agisse des positions représentées sur la figure 1.

En déplaçant la 15-ième, puis la 12-ième fiche, nous obtenons la position B'. Mettons ensuite la fiche n° 1 à la place qu'elle occupe dans A au moyen de la transposition (1, 10) ; effectuons ensuite les transpositions (2, 10), (3, 4), (4, 5), (5, 9), (6, 10), (7, 10), (9, 11) (10, 11), (11, 15). Les fiches 8, 12, 13, 14, 15 restent à leur place et n'interviennent pas dans les transpositions.

Enonçons maintenant sans démonstration le théorème dont nous avons besoin.

● THÉORÈME. *Si la position B peut être transformée en la position A à l'aide de déplacements et d'un nombre pair de transpositions, on peut alors aussi passer de B à A au moyen des seuls déplacements. Si on passe de B à A à l'aide de déplacements et d'un nombre impair de transpositions, alors le passage de B à A à l'aide des seuls déplacements est impossible.*

On voit maintenant clairement comment construire un algorithme pour le problème posé : passons de la position A à la position B selon le procédé indiqué ci-dessus, à l'aide

de déplacements et de transpositions (pas plus de $15 + 15 = 30$ déplacements) ; si l'on a effectué un nombre pair de transpositions, la réponse à la question posée est positive, sinon elle est négative. Ainsi, dans notre exemple le nombre de transpositions est pair, donc on peut effectivement passer de B à A.

2. *Recherche d'un chemin dans un labyrinthe.* La mythologie grecque raconte l'histoire du héros légendaire Thésée qui osa pénétrer dans le labyrinthe pour trouver et tuer le monstrueux Minotaure. Ariane l'aida à sortir du labyrinthe en lui donnant un fil dont elle tenait elle-même une extrémité. A mesure que Thésée s'enfonçait dans le labyrinthe, il déroulait le fil ; il parvint ensuite à trouver la sortie en l'enroulant à nouveau.

C'est à cette vieille légende que fait penser le jouet mécanique relativement récent de la « Souris dans le labyrinthe », mis au point par l'ingénieur et mathématicien américain Claude Shannon. En un point déterminé d'un labyrinthe spécialement conçu, on place un objet figurant un morceau de lard et en un autre point on place une « souris ». La souris se met à errer dans le labyrinthe en zigzags jusqu'à ce qu'elle trouve le « lard ». Si on fait ensuite partir la « souris » du même point, alors elle court tout droit vers sa « nourriture », sans plus faire de détours. Nous considérerons ici un problème analogue de recherche d'un chemin dans un labyrinthe (plus exactement une série de ces problèmes du même type) et décrirons l'algorithme selon lequel doivent s'effectuer les recherches pour atteindre l'objectif indiqué dans le problème.

Représentons le labyrinthe sous forme d'un système fini de carrefours desquels partent des couloirs ; chaque couloir relie deux carrefours (de tels carrefours seront dits contigus), mais on n'exclut pas la possibilité de carrefours dont on ne puisse sortir que par un seul couloir (de tels carrefours seront appelés impasses). On peut géométriquement représenter le labyrinthe sous la forme d'un système

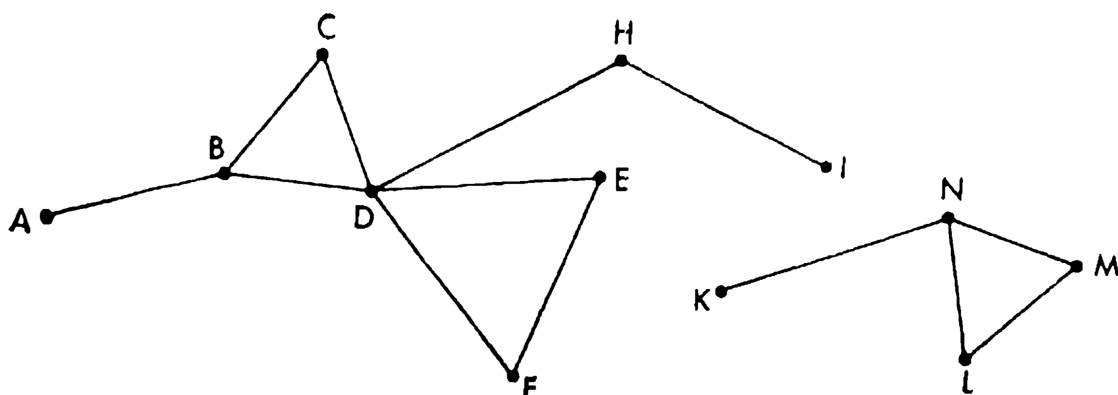


Fig. 2

de points A, B, C... (qui représentent les carrefours) et d'un ensemble de segments AB, BC, . . . (qui représentent les couloirs), et qui relient certains couples de points (fig. 2).

On dira qu'on peut atteindre le carrefour Y à partir du carrefour X s'il existe un chemin allant de X à Y et qui passe par des couloirs et des carrefours intermédiaires. Plus exactement, cela signifie que ou bien X et Y sont des carrefours contigus, ou bien qu'il existe une suite de carrefours X_1, X_2, X_3, . . ., X_n tels que X et X_1, X_1 et X_2, X_2 et X_3, etc., enfin X_n et Y, soient contigus. Ainsi, par exemple, sur la figure 2, on peut atteindre le carrefour H à partir de l'impasse A en empruntant le chemin AB, BC, CD, DE, EF, FD, DH, tandis qu'on ne peut pas atteindre le carrefour K à partir de A. De plus, si on peut atteindre Y à partir de X, on peut toujours l'atteindre par un *chemin simple*, c'est-à-dire par un chemin qui ne passe qu'une seule fois par le même carrefour (et *a fortiori* par le même couloir). Dans l'exemple précédent, le chemin n'était pas simple, mais si l'on supprime le détour DE, EF, FD, on obtient un chemin simple AB, BC, CD, DH.

Supposons que le Minotaure se trouve à l'un des carrefours du labyrinthe (notons-le M), et que Thésée parte à sa recherche à partir du carrefour A, où l'attend Ariane. Celui-ci doit résoudre le problème suivant : savoir si l'on peut ou non

atteindre M à partir de A *). Si oui, il faut l'atteindre en suivant n'importe quel chemin, mais revenir vers Ariane par un chemin simple. Sinon, il faut revenir auprès d'Ariane.

Il peut y avoir un nombre infini de labyrinthes différents, et la disposition mutuelle des carrefours A et M peut aussi varier à l'intérieur d'un labyrinthe donné. Comme Thésée ne sait rien à l'avance sur la disposition du labyrinthe ni sur l'endroit où se trouve le Minotaure, il faut penser la résolution d'un tel problème sous forme d'une méthode générale de recherche qui soit adaptée à n'importe quel labyrinthe et à n'importe quelle disposition des carrefours A et M. En d'autres termes, il faut penser la solution sous forme d'un algorithme qui puisse résoudre n'importe lequel des problèmes du type donné.

Pour construire un tel algorithme, nous adopterons une méthode spéciale de recherche. Selon cette méthode, il faut, à chaque étape du processus de recherche, distinguer entre les couloirs qui n'ont pas encore été empruntés par Thésée (supposons-les verts), ceux qu'il n'a empruntés qu'une seule fois (jaunes) et ceux qu'il a empruntés deux fois (rouges). Ensuite, lorsqu'il se trouve à un carrefour quelconque, Thésée peut aller vers un carrefour contigu en empruntant l'un des deux chemins suivants :

1. *Déroulement du fil.* Passage du carrefour donné au carrefour contigu par n'importe quel couloir vert. (Le fil d'Ariane est alors déroulé le long de ce couloir, qui est considéré comme jaune après le passage de Thésée.)

2. *Enroulement du fil.* Retour du carrefour donné vers le carrefour contigu par le dernier couloir jaune parcouru. Le fil d'Ariane, qui s'est auparavant déroulé le long de ce couloir, s'enroule à nouveau, et ce couloir devient rouge.

On suppose que Thésée dispose de repères qui lui permettent de distinguer par la suite les couloirs rouges des

*) Il est naturel de supposer que $M \neq A$.

verts; on peut distinguer les jaunes, car le fil d'Ariane est déroulé le long de ces couloirs. Le choix d'un chemin ou d'un autre dépend du tableau qui s'offre aux yeux de Thésée au carrefour où il se trouve au moment donné. Ce tableau peut être caractérisé par un ou plusieurs des signes suivants:

1. *Minotaure*. Au carrefour donné on aperçoit le Minotaure.

2. *Détour*. Par le carrefour donné passe déjà le fil d'Ariane; en d'autres termes, deux couloirs jaunes au moins partent de ce carrefour.

3. *Rue verte*. Du carrefour donné part un couloir vert au moins.

4. *Ariane*. Au carrefour donné se trouve Ariane.

5. *Cinquième cas*. Aucun des signes précédents.

On peut maintenant procéder aux recherches suivant le schéma:

Signe	Déplacement
1. *Minotaure*	Arrêt
2. *Détour*	Enroulement du fil
3. *Rue verte*	Déroulement du fil
4. *Ariane*	Arrêt
5. *Cinquième cas*	Enroulement du fil

Lorsqu'il se trouve à l'un quelconque des carrefours, Thésée poursuit son chemin de la façon suivante: il passe en revue les signes de la colonne de gauche du schéma pour savoir dans quel cas il se trouve; dès qu'il a observé le premier de ces signes, il effectue le déplacement (ou l'arrêt) correspondant de la colonne de droite (sans poursuivre le passage en revue des autres signes). Les déplacements s'effectuent jusqu'à ce qu'on rencontre un arrêt.

Le fait que cette méthode soit adaptée au problème résulte immédiatement des trois assertions suivantes:

1. Quelle que soit la position mutuelle des carrefours A et M dans le labyrinthe, on rencontre obligatoirement un arrêt, après un nombre fini de déplacements, soit au carrefour du Minotaure, soit au carrefour d'Ariane.

2. Si l'arrêt a lieu au carrefour du Minotaure, celui-ci est trouvé. De plus, le fil d'Ariane est dans ce cas déroulé le long d'un chemin simple qui va de A à M; en enroulant le fil, Thésée peut maintenant revenir au carrefour d'Ariane par le même chemin.
3. Si l'arrêt a lieu au carrefour d'Ariane, le Minotaure n'a pas été atteint.

Avant de démontrer ces assertions, montrons sur deux exemples comment fonctionne la méthode proposée.

● EXEMPLE 1. Supposons que la recherche du Minotaure, qui se trouve en F, s'effectue à partir du carrefour A du labyrinthe (fig. 2). Il est commode de représenter le processus de recherche qui correspond à notre méthode par le schéma (c'est seulement l'un des schémas possibles, car le choix d'un couloir vert est arbitraire) représenté sur le tableau 1.

Tableau 1

Numéro du déplacement effectué (dans l'ordre)	Signe qui guide Thésée	Déplacement	Couloir emprunté	Couleur du couloir après le passage de Thésée
1	*Rue verte*	Déroulement	AB	Jaune
2	»	»	BC	»
3	»	»	CD	»
4	»	»	DH	»
5	»	»	HJ	»
6	*Cinquième cas*	Enroulement	JH	Rouge
7	»	»	HD	»
8	*Rue verte*	Déroulement	DB	Jaune
9	*Détour*	Enroulement	BD	Rouge
10	*Rue verte*	Déroulement	DF	Jaune
11	*Minotaure*	Arrêt	....	....

On voit que, dans le cas présent, le Minotaure est atteint. Après avoir choisi dans l'avant-dernière colonne ceux des couloirs qui sont restés jaunes (et compte tenu des indications de la dernière colonne), on obtient le chemin simple suivant qui va de A à F:

$$AB,\ BC,\ CD,\ DF.$$

● EXEMPLE 2. Si les recherches partent du carrefour K, on peut alors représenter le processus de ces recherches par le schéma de la figure 2.

On voit que, dans ce cas, le Minotaure n'est pas atteint.

Venons-en maintenant à la démonstration des assertions 1-3.

Démonstration de l'assertion 1. Montrons d'abord, par une récurrence portant sur le nombre de déplacements de Thésée, qu'à chaque étape du processus de recherche se présente l'alternative suivante (i.e. a lieu l'un des deux cas suivants qui s'excluent mutuellement):

Tableau 2

Numéro du déplacement effectué (dans l'ordre)	Signe qui guide Thésée	Déplacement	Couloir emprunté	Couleur du couloir après le passage de Thésée
1	*Rue verte*	Déroulement	KN	Jaune
2	»	»	NL	»
3	»	»	LM	»
4	»	»	MN	»
5	*Détour*	Enroulement	NM	Rouge
6	*Cinquième cas*	»	ML	»
7	»	»	LN	»
8	»	»	NK	»
9	*Ariane*	Arrêt	...	....

a) Il n'y a pas dans le labyrinthe de couloirs jaunes; Thésée se trouve au carrefour A (Ariane).

b) Il y a des couloirs jaunes dans le labyrinthe et, si on les considère dans l'ordre où les a parcourus Thésée, ils constituent un chemin allant de A au carrefour où se tient Thésée.

Il s'avérera de plus que Thésée n'emprunte jamais de couloir rouge.

Les cas décrits sont évidents au début du processus, lorsque Thésée se trouve encore au carrefour A et n'a encore emprunté aucun couloir (tous les couloirs sont verts). Supposons maintenant que les alternatives indiquées soient vraies après le $(n - 1)$-ième déplacement et montrons qu'elles sont encore vraies après le n-ième (dans le seul cas, naturellement, où le $(n - 1)$-ième déplacement n'a pas conduit à un arrêt).

Supposons que l'on ait le cas a après le $(n - 1)$-ième déplacement. Alors le déplacement suivant peut se faire suivant un couloir vert allant de A à un carrefour contigu K (après le n-ième déplacement a lieu le cas b avec un seul couloir jaune AK), ou bien il consiste en un arrêt au carrefour A (après le n-ième déplacement, on conserve le cas a).

Supposons maintenant qu'après le $(n - 1)$-ième déplacement, on ait le cas b avec s couloirs jaunes qui forment un chemin AA_1, A_1A_2, . . ., $A_{s-1}K$. Suivant le signe qui guide Thésée lorsqu'il choisit le n-ième déplacement, on a les possibilités suivantes:

1. *Minotaure*. Il se produit un arrêt au carrefour K, avec conservation des couloirs qui étaient antérieurement jaunes (cas b après le n-ième déplacement).

2. *Détour*. Thésée enroule le fil, i.e. emprunte le couloir jaune KA_{s-1}, qui devient maintenant rouge. Le chemin jaune est réduit à un seul couloir; si le nombre s de couloirs du chemin précédent était supérieur à 1, alors, après le n-ième déplacement, on sera en présence du cas b avec $(s - 1)$ couloirs jaunes; si $s = 1$, on aura le cas a.

3. *Rue verte.* Thésée déroule le fil, i.e. avance dans un couloir vert qui devient maintenant jaune. On est en présence du cas b avec $(s+1)$ couloirs jaunes.

4. *Ariane.* Ce signe ne guidera pas Thésée, car s'il revient au carrefour d'Ariane, en suivant le chemin jaune

$$AA_1,\ A_1A_2,\ \ldots,\ A_{s-1}K$$

(i.e. si $K = A$), alors, conformément à la méthode de recherche adoptée, c'est le signe *détour* qui doit le guider.

5. En l'absence des quatre premiers signes, le fil est enroulé, ce qui amène aussi, comme dans le cas d'un détour, au cas a (pour $s = 1$) ou au cas b (pour $s > 1$).

Ainsi, on a complètement établi l'alternative stipulée plus haut ; on a également mis en évidence le fait que Thésée ne passe pas plus de deux fois dans le même couloir (et qu'il n'emprunte jamais un couloir rouge). En outre, dans la mesure où le nombre de couloirs du labyrinthe est fini, la suite des déplacements doit aussi être finie ; mais ces déplacements ne peuvent s'arrêter que si Thésée parvient au carrefour du Minotaure ou à celui d'Ariane.

Démonstration de l'assertion 2. Dans le cas où Thésée s'arrête au carrefour du Minotaure, il est évident que ce dernier peut être atteint ; de plus, le fil d'Ariane est déroulé suivant un chemin jaune dont l'existence a été établie plus haut. L'absence de détours s'explique par ce que, à chaque fois qu'il en découvre un sur son chemin, Thésée fait marche arrière et le supprime du même coup.

Démonstration de l'assertion 3. Remarquons que dans le cas d'un arrêt au carrefour d'Ariane

1) chaque couloir du labyrinthe est ou bien parcouru deux fois (couloir rouge), ou bien aucune (couloir vert) ; en d'autres termes, le fil est complètement enroulé (il n'y a pas de couloirs jaunes dans le labyrinthe) *) ;

---

*) La démonstration de ces faits est laissée au lecteur.

2) tous les couloirs qui partent de A sont rouges car, d'après nos conventions, le signe *Ariane* n'entre en ligne de compte que lorsque les trois signes qui le précèdent dans le schéma, y compris le signe *rue verte*, n'ont pas lieu.

Démontrons maintenant par l'absurde que le Minotaure est atteint et soit AA_1, A_1A_2, . . ., A_nM le chemin qui va de A à M. Le premier couloir de ce chemin est rouge, car il part de A, le dernier est vert, car Thésée n'est pas parvenu à trouver le Minotaure. Soit A_iA_{i+1} le premier couloir vert de cette suite. Ainsi, des couloirs verts et rouges débouchent sur le carrefour A_i. Considérons maintenant le dernier passage de Thésée au carrefour A_i; il est évident que ce passage s'est fait par l'un des couloirs rouges débouchant sur A_i et que le fil a pu seulement être enroulé lors de ce passage, puisque ce n'est pas la première fois que ce couloir a été emprunté. Il est donc motivé ou bien par l'existence d'un détour, ou bien par le cinquième cas, i.e. par l'absence des quatre premiers signes. Mais la dernière situation ne peut avoir lieu, puisque de A_i part le couloir vert A_iA_{i+1}. Il faut donc admettre que le dernier passage par le carrefour A_i est lié à l'existence d'un détour. Mais ceci conduit immédiatement à une contradiction, ce qui achève la démonstration de l'assertion 3. En effet, s'il y avait un détour en A_i, cela signifierait que deux chemins jaunes au moins partent de ce carrefour; lors du déplacement suivant, Thésée aurait rendu l'un d'entre eux rouge; mais le couloir jaune restant doit obligatoirement être parcouru à nouveau, car il ne doit pas rester de couloirs jaunes; donc Thésée passera encore une fois par le carrefour A_i. Ceci contredit l'hypothèse selon laquelle on considère le dernier des passages de Thésée par le carrefour A_i.

Faisons maintenant une remarque. Dans le programme que nous avons proposé pour rechercher un chemin se glisse une certaine part de hasard. En effet, dans la condition *rue verte*, le déplacement suivant n'est pas déterminé de façon unique, puisque du carrefour donné peuvent partir

plusieurs couloirs verts; or, notre programme n'indique pas lequel d'entre eux il faut choisir; plus exactement, il laisse la possibilité de choisir arbitrairement l'un d'entre eux au hasard. On perd par là même la propriété de déterminisme, qui est, comme nous l'avons dit au paragraphe précédent, propre à tous les algorithmes. On pourrait cependant supprimer facilement cette part de hasard (et du même coup transformer notre programme en algorithme) ne serait-ce qu'en convenant d'une règle permettant de choisir l'un des couloirs verts lorsqu'il en existe plusieurs. Par exemple, lorsqu'il débouche sur un carrefour, Thésée commence à en faire le tour dans le sens des aiguilles d'une montre jusqu'à ce qu'il rencontre le premier couloir vert qu'il emprunte ensuite; on pourrait adopter n'importe quelle autre convention. L'étude des programmes où l'on examine à l'avance les actes d'un choix aléatoire présente un intérêt théorique et pratique considérable, surtout dans la théorie moderne des jeux. Cependant nous n'examinerons pas cette question et nous nous contenterons d'étudier des processus strictement déterminés.

En conclusion, il est utile de comparer les deux problèmes de ce paragraphe et de mettre en évidence une certaine corrélation entre eux. Malgré leur différence apparente, ces deux problèmes sont au fond très voisins. Plus exactement, le premier est un cas particulier du second. En effet, si l'on associe un carrefour à chacune des positions du jeu du « 15 » et un couloir à chaque passage d'une position à une position contiguë, reliant les carrefours correspondants, alors le premier problème de ce paragraphe devient le problème de recherche d'un chemin dans un labyrinthe d'un type particulier, comportant 16! carrefours, chacun d'entre eux communiquant par un couloir avec deux, trois ou quatre carrefours contigus. Mais dans ce cas, l'algorithme que nous avons proposé pour rechercher un chemin peut s'appliquer au jeu du « 15 ». Il est facile de comprendre qu'il représente simplement une variante perfectionnée de l'algorithme de traite-

ment ; l'amélioration consiste en ce que, lorsqu'on trie les positions et on se déplace, on ne tolère pas plus de deux répétitions.

Pour un labyrinthe d'un type particulier (correspondant au jeu du « 15 »), on a réussi à trouver un algorithme plus simple.

De plus, il est tout naturel de supposer que, dans le cas général où l'algorithme doit convenir à n'importe quel labyrinthe, il ne peut représenter rien d'autre qu'un certain type de tri. C'est pourquoi on peut difficilement espérer construire un algorithme plus simple que celui que nous avons proposé.

§ 3. PROBLÈME DES MOTS

Le problème des mots est une généralisation poussée du jeu du « 15 » et des recherches de Thésée. Si le jeu du « 15 » se ramène à des recherches effectuées dans un labyrinthe fini, fixé à l'avance, et si les recherches de Thésée peuvent avoir lieu dans n'importe quel labyrinthe fini, le problème des mots, par contre, est, dans un certain sens, un problème de recherche dans un labyrinthe infini. Le problème des mots a pris naissance dans certains secteurs particuliers de l'algèbre moderne, appelés *théorie des systèmes associatifs* et *théorie des groupes* ; son importance dépasse toutefois le cadre de ces théories spéciales. Les grands mathématiciens soviétiques André Markov et Piotr Novikov, ainsi que leurs élèves, ont étudié avec succès différentes variantes de ce problème.

Introduisons d'abord quelques notions préliminaires :

Nous appellerons *alphabet* tout système fini de signes deux à deux distincts, appelés *lettres* de cet alphabet. On peut ainsi, par exemple, considérer l'alphabet $\{\alpha, ц, z, ?\}$ constitué de la lettre grecque α, de la lettre russe ц, de la

lettre latine z et du point d'interrogation. Dans un tel alphabet, nous appellerons *mot* toute suite de lettres de cet alphabet. Ainsi *abaa* et *bbac* sont des mots dans l'alphabet $\{a, b, c\}$. Si le mot L est une partie du mot M, nous parlerons d'*inclusion* du mot L dans le mot M. Par exemple, dans le mot *abc bc bab*, on a deux inclusions du mot *bcb*: l'une en commençant par la seconde lettre, l'autre en commençant par la quatrième. Nous étudierons les transformations de certains mots dans d'autres, effectuées au moyen de certaines substitutions permises, qui sont données sous la forme

$$P - Q \quad \text{ou} \quad P \to Q,$$

où P et Q sont deux mots du même alphabet.

On peut faire subir une *substitution orientée* $P \to Q$ à un mot R dans le cas où le premier membre P est au moins une fois inclus dans R ; la substitution consiste à remplacer l'une quelconque de ces inclusions par le second membre correspondant Q. Quant à l'application d'une *substitution non orientée* $P - Q$, elle consiste aussi bien à remplacer une inclusion du premier membre par le second qu'une inclusion du second membre par le premier. A partir de maintenant, nous considérerons surtout des substitutions non orientées, que nous appellerons simplement substitutions, s'il n'y a pas de confusion possible.

● EXEMPLE 1. La substitution *ab* — *bcb* est applicable de quatre manières différentes au mot *abcbcbab* ; le remplacement de chacune des deux inclusions *bcb* donne les mots :

$$a\,\underline{ab}\,cb\,ab, \quad abc\,\underline{ab}\,ab,$$

et le remplacement de chacune des deux inclusions *ab* donne les mots :

$$\underline{bcb}\,cbc\,bab, \quad abc\,bcb\,\underline{bcb}.$$

Cette substitution ne s'applique pas au mot *bacb*.

Convenons d'appeler *calcul associatif* l'ensemble de tous les mots d'un alphabet muni d'un système fini de substitutions permises.

Pour se donner un calcul associatif, il suffit d'indiquer l'alphabet et le système de substitutions correspondants.

Si on peut transformer le mot R en le mot S par une seule application d'une substitution permise, alors S peut être aussi transformé en R par le même procédé ; dans ce cas, les mots R et S seront dits *voisins*. Une suite de mots

$$R_1, R_2, \ldots, \quad R_{n-1}, R_n$$

tels que R_1 et R_2, R_2 et R_3, . . ., R_{n-1} et R_n soient voisins sera appelée *chaîne déductive*, allant de R_1 à R_n. S'il existe une chaîne déductive allant du mot R au mot S, il existe évidemment aussi une chaîne déductive qui va de S à R ; dans ce cas, R et S seront dits *équivalents*, et nous noterons : $R \sim S$. Il est clair que si $S \sim R$ et $R \sim T$, alors $S \sim T$. Nous aurons encore besoin par la suite du théorème suivant :

● THÉORÈME. *Soit $P \sim Q$; alors, si l'on a une inclusion de P dans un mot R, le mot obtenu par substitution de Q à cette inclusion est équivalent à R.*

● DÉMONSTRATION. Dans les conditions du théorème, il est commode de représenter le mot R sous la forme SPT, où S est la partie du mot R qui précède l'inclusion de P et T celle qui la suit, et le mot transformé sous la forme SQT. D'autre part, puisque P est équivalent à Q, il existe une chaîne déductive

$$P, P_1, P_2, \ldots, P_m, Q.$$

Mais, dans ce cas, on voit facilement que la suite de mots

$$SPT, SP_1T, SP_2T, \ldots, SP_mT, SQT$$

est une chaîne déductive allant de SPT (i.e. R) à SQT (i.e. au mot transformé). Le théorème est démontré.

● EXEMPLE 2. Considérons le calcul associatif étudié par C. Tséitine.

Alphabet :

$$\{a,\ b,\ c,\ d,\ e\}.$$

Système de transformations permises :

ac — *ca*
ad — *da*
bc — *cb*
bd — *db*
abac — *abacc*
eca — *ae*
edb — *be*

Dans ce calcul, on ne peut appliquer au mot *abcde* que la troisième substitution, et ce mot n'a qu'un mot voisin *acbde*. On a d'autre part l'équivalence : *abcde* ∼ *cadedb*, comme il résulte de la chaîne déductive suivante :

$$a\underline{bc}de,\quad \underline{ac}bde,\quad ca\underline{bd}e,\quad cad\underline{be},\quad cadedb.$$

Si par contre on prend le mot *aaabb*, on ne peut lui appliquer aucune des substitutions, et il n'a par suite aucun mot voisin ; il n'existe *a fortiori* aucun mot différant de *aaabb* qui lui soit équivalent.

Pour chaque calcul associatif se pose le problème spécifique d'*équivalence des mots*. Il consiste en ceci :

Pour deux mots quelconques d'un calcul donné, trouver s'ils sont ou non équivalents.

Dans la mesure où il y a une infinité de mots distincts dans tout calcul, on est en fait ici en présence d'une série infinie de problèmes du même type ; la résolution est donnée sous la forme d'un algorithme qui décide de l'équivalence ou de la non-équivalence de tout couple de mots.

On peut avoir l'impression que le problème des mots est un casse-tête artificiel et que sa mise au point ne présente par suite aucun intérêt théorique ou pratique particulier. En fait, c'est tout le contraire; on peut montrer que ce problème est parfaitement naturel et a une importance théorique et pratique considérable qui justifie pleinement les efforts déployés pour construire l'algorithme correspondant. Toutefois, à ce stade de l'exposé, nous nous garderons d'examiner le fond de la question et passerons à l'étude de certains faits concrets.

Tout d'abord, montrons le lien entre le problème d'équivalence des mots et le problème de Thésée. Si pour chaque mot on construit un « carrefour » et pour chaque couple de mots voisins un couloir reliant les carrefours correspondants, le calcul associatif nous apparaîtra alors sous la forme d'un labyrinthe ayant un nombre infini de carrefours et de couloirs; de chaque carrefour part seulement un nombre fini de couloirs (il peut y avoir aussi, bien sûr, des carrefours d'où ne part aucun couloir: cf. le mot *aaabb* de l'exemple 2). D'autre part, une chaîne déductive qui va d'un mot R à un mot Q représentera un chemin dans le labyrinthe qui va d'un carrefour à un autre; ainsi, à l'équivalence des mots correspondra le fait que l'on puisse atteindre un carrefour à partir d'un autre. Enfin, vu sous cet angle, le problème des mots lui-même devient un problème de recherche d'un chemin dans un labyrinthe.

Pour mieux comprendre l'aspect particulier des difficultés qui se présentent ici, nous allons au préalable étudier un *problème restreint des mots*, qui consiste en ceci:

Etant donné deux mots R et T d'un calcul associatif donné, peut-on passer de l'un à l'autre en appliquant successivement des substitutions permises dont le nombre ne dépasse pas k (où k est un entier naturel arbitraire, fixé à l'avance)?

Si le problème est posé de cette façon, on peut facilement construire un algorithme: on peut utiliser l'algorithme de tri que nous connaissons déjà, et dans lequel on considère

une liste de mots, en commençant par le mot R, puis en écrivant tous les mots voisins de R, les mots voisins de ces derniers, etc., k fois. La réponse à la question posée sera positive ou négative suivant que le mot T figure ou non dans cette liste.

Toutefois, si on revient au problème non restreint des mots, la situation est foncièrement différente. Comme la chaîne allant de R à T (si elle existe) peut être arbitrairement longue, on ne sait pas, en général, à quel moment il faut considérer que le processus de tri est achevé. Supposons par exemple que l'on ait poursuivi le processus de tri jusqu'à 10^{20} = 100 000 000 000 000 000 000, et que l'on dispose déjà de la liste de tous les mots qu'on puisse obtenir de R par applications de substitutions, en nombre total inférieur à 10^{20}. Supposons que l'on n'ait pas rencontré le mot T dans cette liste. Cela nous permet-il de conclure que les mots R et T ne sont pas équivalents? Certainement pas, car il n'est pas exclu que les mots R et T soient équivalents, mais que la chaîne déductive minimale les reliant soit encore plus longue.

Pour obtenir les résultats souhaités, il faut ici renoncer à un tri simple; il faut trouver d'autres idées, fondées sur l'analyse du mécanisme lui-même de transformation des mots les uns dans les autres au moyen des substitutions permises.

Essayons, par exemple, de dire si les mots *abaacd* et *acbdad* sont équivalents dans le calcul de Tséitine (cf. exemple 2). La réponse est négative, comme il résulte des considérations suivantes: dans chacune des substitutions permises de ce calcul, le premier et le second membres contiennent le même nombre d'inclusions de la lettre a (ou ne contiennent pas du tout cette lettre); donc, dans toute chaîne déductive, tous les mots doivent contenir le même nombre d'inclusions de la lettre a. Comme ce nombre est différent dans les deux mots que nous avons considérés, ces deux mots ne sont pas équivalents.

La détermination de tels *invariants déductifs*, i.e. de propriétés qui restent invariantes pour tous les mots d'une même chaîne déductive, permet dans certains cas de trouver les algorithmes cherchés.

● EXEMPLE 3.
Alphabet

$$\{a, b, c, d, e\}.$$

Système de substitutions permises :

$$ab - ba; \quad ae - ea; \quad be - eb; \quad de - ed;$$
$$ac - ca; \quad bc - cb; \quad cd - dc;$$
$$ad - da; \quad bd - db; \quad ce - ec.$$

Ces substitutions permises ne changent pas le nombre des inclusions de chaque lettre dans un mot ; seul change l'ordre des lettres dans le mot. On voit immédiatement que deux mots sont équivalents si et seulement si le nombre d'inclusions de chaque lettre est le même dans les deux mots. C'est pourquoi l'algorithme qui permet de déterminer l'équivalence est très simple et se ramène au calcul du nombre d'inclusions des lettres dans chacun des mots et à une comparaison de ces nombres.

Plus loin, nous traiterons en détail un exemple bien plus compliqué, mais auparavant nous allons définir les généralisations suivantes des notions de « mot » et de « substitution permise ». En effet, outre les mots habituels, nous considérerons également dans l'alphabet donné le *mot vide*, qui ne contient aucune lettre, et que nous désignerons par Λ. De plus, nous admettrons une substitution du type

$$P - \Lambda.$$

Le remplacement du premier membre par le mot vide signifie simplement que l'on supprime du mot transformé

l'inclusion du mot P. Le remplacement du second membre par le premier signifie qu'entre deux lettres quelconques du mot transformé, ou devant lui, ou derrière, on ajoute le mot P.

● EXEMPLE 4. Donnons-nous un calcul associatif dans l'alphabet $\{a, b, c\}$, avec le système de substitutions permises :

$$b - acc$$
$$ca - accc$$
$$aa - \Lambda$$
$$bb - \Lambda$$
$$cccc - \Lambda$$

On demande de trouver l'algorithme pour le problème d'équivalence des mots dans ce calcul.

Construisons un algorithme auxiliaire, ou algorithme de réduction, qui donne pour chaque mot un mot équivalent d'un type particulier, ou mot réduit. Pour cela, considérons le système ordonné des substitutions orientées

$$b \rightarrow acc$$
$$ca \rightarrow accc$$
$$aa \rightarrow \Lambda$$
$$cccc \rightarrow \Lambda$$

et supposons que, si on applique à un mot quelconque R, l'algorithme fonctionne ainsi : on prend la première des substitutions orientées de cette liste qui soit applicable au mot R et, lorsqu'il existe plusieurs inclusions de son premier membre dans le mot R, on applique la substitution à l'inclusion située le plus à gauche dans R ; on obtient ainsi un mot R' ; pour R', on prend à nouveau dans la liste la première des substitutions qui lui soit applicable, et on l'applique à l'inclusion de son premier membre située le

plus à gauche dans R' ; si, après un nombre fini de telles opérations, on obtient un mot S auquel aucune des substitutions de la liste n'est plus applicable, on dit que l'algorithme est applicable au mot R et qu'il le transforme en un mot S *).

Montrons que, quel que soit le mot R, l'algorithme de réduction est applicable à R et le transforme en l'un des huit mots suivants (mots réduits) :

$$\wedge, c, cc, ccc, a, ac, acc, accc.$$

En effet, si dans le mot R figurent des inclusions de la lettre b, on appliquera tout d'abord la première substitution, en supprimant b et en le remplaçant par acc, jusqu'à ce qu'il ne reste plus de lettre b. Ensuite, on applique la seconde substitution, en transformant une lettre a qui précède une lettre c, jusqu'à ce qu'il ne reste plus aucune lettre a suivant immédiatement une lettre c ; en d'autres termes, jusqu'à ce que tous les a précèdent tous les c. Enfin, on commence à supprimer tous les couples formés de deux a voisines et tous les quadruplets de c voisines, jusqu'à ce qu'apparaisse un mot ne comportant pas plus d'un a et de trois c ; il peut y avoir en tout huit de ces mots ; ce sont justement ceux que nous avons énumérés plus haut.

Il est évident que chaque mot est équivalent au mot que l'on obtient en le réduisant ; c'est pourquoi deux mots sont équivalents si et seulement s'il leur correspond un même mot réduit, ou deux mots réduits équivalents. Nous démontrerons plus bas que les huit mots réduits sont deux à deux non équivalents. De là il résultera que deux mots sont équivalents si et seulement s'il leur correspond un même mot

*) Les algorithmes de tri des mots dans un alphabet, qui sont donnés de cette façon par un système ordonné de substitutions orientées, sont appelés normaux. La théorie des algorithmes normaux présente un grand intérêt théorique et pratique; elle a été mise au point par le membre de l'Académie des Sciences de l'U.R.S.S. A. Markov et exposée dans son livre « Théorie des algorithmes ».

réduit. On construira de plus un algorithme pour le problème des mots posé. Il consistera à appliquer l'algorithme de réduction à chacun des deux mots considérés et à comparer ensuite les mots réduits obtenus.

Soient par exemple les deux mots *cacb* et *bb*. On obtient les mots réduits:

1) *cacb*, *cacacc*, *accccacc*, *acccaccccc*, *accaccccccccc*, *acacccccccccccc*, *aacccccccccccccc*, *cccccccccccccc*, *cccccccccc*, *cccccc*, *cc*;

2) *bb*, *accb*, *accacc*, *acaccccc*, *aacccccccc*, *cccccccc*, *cccc*, Λ.

Conclusion: les mots *cacb* et *bb* ne sont pas équivalents, car on a obtenu deux mots réduits différents: *cc* et Λ.

Démonstration de la non-équivalence des huit mots réduits, pris deux à deux. Remarquons tout d'abord que, étant donné une chaîne déductive allant d'un mot R ne contenant pas la lettre b à un mot S qui ne la contient pas non plus, on peut en extraire une chaîne déductive allant de R à S telle que les mots intermédiaires de cette chaîne ne contiennent pas la lettre b. En effet, si dans tous les mots de la chaîne déductive donnée on remplace chaque inclusion de la lettre b par le mot *acc*, on obtient une suite de mots dans laquelle deux mots consécutifs sont voisins (au sens du calcul) ou simplement identiques. Si l'on supprime maintenant les mots qui se répètent (parmi les mots consécutifs), on obtient alors la chaîne déductive cherchée. La substitution *b—acc* n'intervient pas dans les chaînes déductives de ce type.

D'autre part, dans chacune des substitutions permises restantes, les inclusions de la lettre a dans le premier et dans le second membre sont en nombre ou bien simultanément pair, ou bien simultanément impair. Il en va de même pour la lettre c. Donc la parité du nombre d'inclusions de la lettre a (ou de la lettre c) est un invariant déductif pour les chaînes déductives du type indiqué plus haut. Il en résulte immédiatement qu'aucun des quatre mots réduits contenant une fois l'inclusion de la lettre a n'est équivalent à l'un

des quatre mots réduits qui ne contiennent pas cette inclusion. Exactement de la même façon, aucun des mots réduits contenant une ou trois inclusions de la lettre *c* n'est équivalent à un mot en contenant deux ou n'en contenant pas. C'est pourquoi il reste encore à s'assurer de la non-équivalence des couples suivants de mots:

$$\Lambda, \ cc; \ c, ccc; \ a, acc; \ ac, accc.$$

Si l'on avait équivalence d'au moins l'un des trois premiers couples, alors on aurait aussi équivalence du quatrième, comme il ressort du théorème de ce paragraphe. Donc il suffit d'établir la non-équivalence du couple *ac*, *accc*; passons maintenant à cette démonstration.

Définissons les termes suivants.

On appelle *indice de l'inclusion de la lettre a* dans le mot *R* le nombre de toutes les inclusions de la lettre *c* qu'on rencontre à droite de cette inclusion de la lettre *a*. On appelle *indice du mot R* la somme des indices de toutes les inclusions de la lettre *a* *). Aucune des deux substitutions $aa-\Lambda$ et $cccc-\Lambda$ ne change la parité de l'indice du mot. En effet, si l'on substitue *aa* au mot vide, l'indice du mot est augmenté de la somme des indices, égaux entre eux, des deux inclusions de la lettre *a*, i.e. d'un nombre pair; lorsqu'on remplace *aa* par le mot vide, l'indice du mot est diminué d'un nombre pair.

Lorsqu'on effectue la substitution *cccc*, les indices de certaines inclusions de *a* augmentent de quatre, les indices des autres ne changent pas; l'indice total du mot augmente d'un nombre pair. Les conclusions sont les mêmes lorsqu'on supprime *cccc*. La substitution $b-acc$ ne change évidemment pas la parité de l'indice du mot.

---

*) Par exemple, dans le mot *acbca*, la première inclusion à partir de la gauche de la lettre *a* a l'indice 0 (il n'y a pas d'inclusion de la lettre *c* située plus à gauche); la seconde inclusion de la lettre *a* a l'indice 2. L'indice du mot est 2.

Montrons enfin que la substitution *ca—accc* change la parité de l'indice du mot. Pour cela, comparons les mots

PcaQ et *PacccQ* ;

l'indice de chacune des inclusions de *a* dans la partie *P* du mot *R* varie exactement de 2, tandis que l'indice de l'inclusion de *a* dans *Q* est inchangé. L'indice de l'inclusion unique de *a* entre *P* et *Q* varie exactement de 3. En tout, l'indice du mot varie d'un nombre impair.

Les mots *ac* et *accc* ont tous deux des indices de même parité ; c'est pourquoi, s'ils sont équivalents, alors, dans la chaîne déductive correspondante (on peut considérer qu'on ne rencontre pas la lettre *b* dans cette chaîne, cf. ci-dessus) ne peut figurer qu'un nombre pair des substitutions *ca—accc*.

Toutefois une telle hypothèse conduit à une contradiction, comme il ressort du raisonnement suivant. Chaque application de la substitution *ca—accc* change le nombre d'inclusions de la lettre *c* de 2 unités ; c'est pourquoi, si on l'applique un nombre pair de fois, on fait varier le nombre d'inclusions de la lettre *c* d'un nombre multiple de quatre. Evidemment, la substitution $cccc - \Lambda$ fait varier le nombre d'inclusions de la lettre *c* exactement de quatre unités, et la substitution $aa - \Lambda$ ne fait pas varier ce nombre. En faisant le bilan de ce qui a été dit, il faut conclure que si $ac \sim accc$, alors les nombres d'inclusions de la lettre *c* dans ces deux mots diffèrent entre eux d'un multiple de quatre, ce qui n'est pas le cas. Donc, les mots *ac* et *accc* ne sont pas équivalents. Ceci achève de justifier l'algorithme que nous avons proposé.

La solution détaillée du problème des mots que nous avons exposée dans l'exemple 4 pour un calcul associatif concret caractérise à bien des égards les méthodes et les notions qui apparaissent lorsqu'on examine le problème des mots pour d'autres calculs. Il nous reste à éclaircir la signification que renferme le problème des mots et son importance

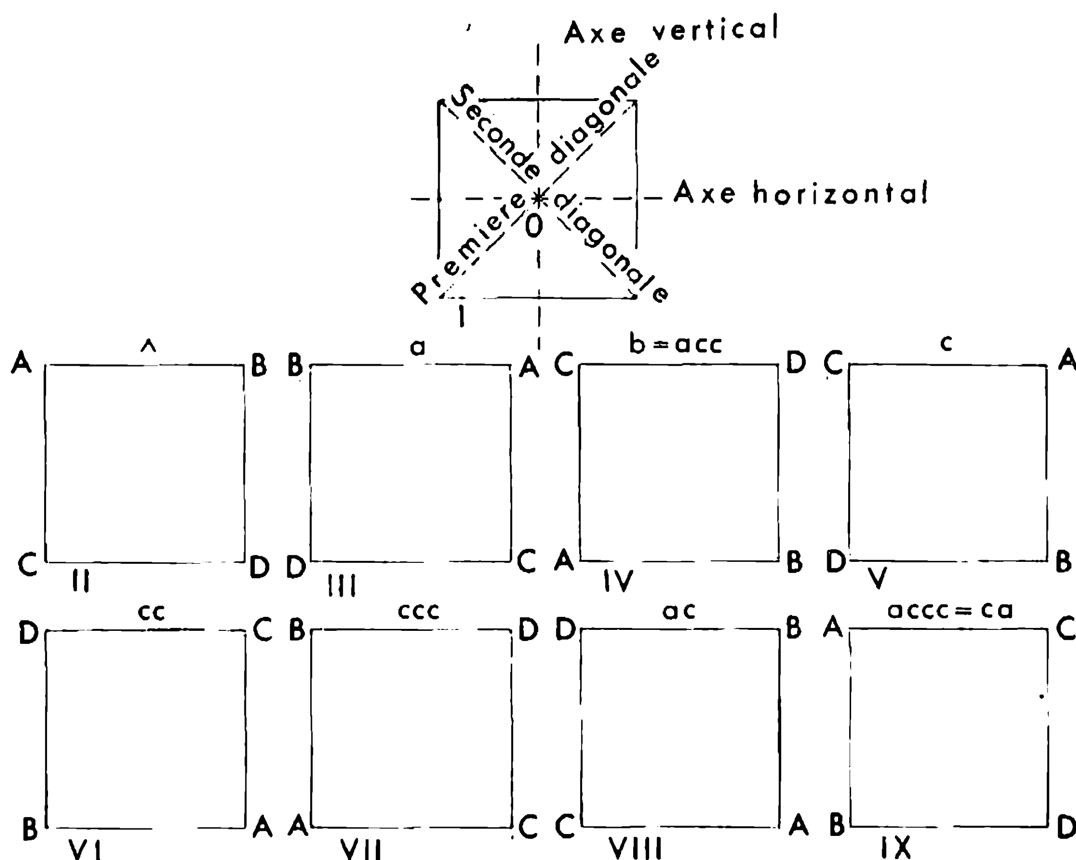


Fig. 3

en algèbre moderne. Il est commode de la faire sur un exemple concret.

Prenons un carré quelconque (fig. 3, I). Considérons les trois *endomorphismes* suivants de ce carré, i.e. les trois transformations suivantes du carré en lui-même:

1) la symétrie (réflexion par un miroir) par rapport à un axe vertical passant par le centre du carré O;

2) la symétrie par rapport à un axe horizontal passant par le centre du carré O;

3) la rotation de 90° autour du centre O dans le sens des aiguilles d'une montre. On appellera ces transformations transformations élémentaires et nous les désignerons par les lettres a, b et c respectivement. Sur la figure 3 (III-IV-V) on indique comment varie la disposition des sommets du

carré (II) lorsqu'on applique chacune des transformations élémentaires.

Remarquons que la réalisation successive de deux ou de plusieurs automorphismes quelconques du carré est encore un endomorphisme du carré. Nous allons dans la suite nous en tenir à la définition généralement admise selon laquelle multiplier deux transformations données (en particulier, deux automorphismes du carré) signifie les appliquer successivement. Convenons en plus de conserver, pour l'opération de multiplication des transformations, la terminologie usuelle ainsi que le terme de *produit* pour la transformation résultante. Ainsi, par exemple, le produit cc représente le résultat de deux rotations successives de 90°, c'est-à-dire une rotation de 180° ; le produit ac, de la symétrie par rapport à l'axe vertical suivi d'une rotation de 90°, équivaut à une symétrie par rapport à la seconde diagonale (voir fig. 3, I). Le produit (ac) (cc) des deux produits précédents équivaut à une symétrie par rapport à la première diagonale (voir fig. 3, I).

Une telle multiplication n'est évidemment pas commutative ; sur la figure 3 (VIII-IX) on a représenté deux dispositions différentes des sommets du carré qui correspondent aux deux automorphismes ac, ca du carré initial (II). Mais elle possède en revanche, tout comme la multiplication arithmétique usuelle, la propriété d'associativité : quelles que soient les transformations p, q, r, on a l'identité $(pq)r = p(qr)$. Grâce à cette propriété, on peut supprimer les parenthèses, ainsi, par exemple, (ac) (cc) et $(((ac)c)c)$ représentent un seul et même automorphisme du carré, à savoir la symétrie par rapport à la seconde diagonale.

Nous allons examiner l'ensemble Ω, formé des transformations élémentaires a, b, c et de tous les automorphismes du carré que l'on peut considérer comme le produit d'un nombre fini (mais arbitraire) de transformations élémentaires. L'associativité d'une telle multiplication permet d'omettre les parenthèses lorsqu'on écrit symboliquement les élé-

ments de Ω et de se limiter à une écriture qui respecte l'ordre des lettres représentant les automorphismes correspondants, telle que *abb*, *cabb*, *accc*, etc. Cela signifie que tout produit prend la forme d'un mot de l'alphabet $\{a, b, c\}$.

De l'associativité de la multiplication s'ensuit également que, si l'on écrit un mot Q à la droite d'un mot P pour obtenir un seul mot PQ, celui-ci représentera alors le produit des automorphismes symbolisés par les mots P et Q respectivement ; ainsi, par exemple, le mot *abccab* représente le produit des automorphismes symbolisés par les mots *abc* et *cab*.

Il est alors clair qu'il existe dans l'alphabet $\{a, b, c\}$ une infinité de mots graphiquement différents (i.e. qui diffèrent par leur écriture) ; néanmoins, des mots graphiquement différents peuvent représenter un seul et même automorphisme de Ω. Il est naturel dans ce cas de considérer ces mots comme égaux et de noter une telle égalité de la façon usuelle. Le lecteur peut vérifier facilement les égalités suivantes :

$$b = acc, \tag{1}$$

$$ca = accc. \tag{2}$$

Pour cela il suffit de comparer les dispositions des sommets du carré correspondant aux transformations des premier et deuxième membres de ces égalités. Ensuite, il est aisé de voir que chacun des mots : *aa*, *bb*, *cccc* définit un seul et même automorphisme, à savoir la transformation dite identique, pour laquelle tous les sommets restent inchangés. Comme une telle transformation ne change rien, il est raisonnable de la désigner aussi par le mot vide Λ. Ainsi, on a les égalités suivantes

$$aa = \Lambda, \tag{3}$$

$$bb = \Lambda, \tag{4}$$

$$cccc = \Lambda. \tag{5}$$

Si l'on compare les égalités (1)-(5) avec les substitutions permises du calcul associatif de l'exemple 4, on obtient la proposition suivante, qui fit le lien entre ce calcul et le système considéré de transformations du carré:

Deux produits d'automorphismes élémentaires du carré définissent la même transformation si et seulement si les mots les représentant dans le calcul de l'exemple 4 sont équivalents.

En effet, des égalités (1)-(5) résultent que toute substitution permise appliquée à un mot S le transforme en un mot égal. Ainsi, par exemple, la substitution *ca-accc* transforme le mot *bcac* en *bacccc*; mais l'associativité de la multiplication permet d'écrire: $bcac = b\,(ca)\,c$ et $bacccc = b\,(aacc)\,c$; les seconds membres sont égaux comme produits de facteurs respectivement égaux, donc les premiers membres le sont également. Ainsi, *deux mots voisins sont égaux.*

Il est maintenant aisé de comprendre que l'équivalence de deux mots du calcul associatif considéré entraîne leur égalité (c'est-à-dire l'identité des automorphismes qu'ils définissent). En effet, si $S \sim T$, deux éléments voisins de la chaîne correspondante sont égaux, d'où $S = T$.

L'affirmation réciproque est vraie aussi: *si des mots sont égaux, alors ils sont équivalents.* En effet, si deux mots sont égaux, les mots réduits correspondants le sont également (ceci découle de l'affirmation directe). On peut de plus vérifier immédiatement que les huit mots réduits définissent des automorphismes deux à deux distincts (voir fig. 3, II-IX où l'on a représenté les dispositions des sommets du carré (fig. 3, II) pour des automorphismes correspondant aux 8 mots réduits). C'est pourquoi si deux mots sont égaux, il leur correspond alors le même mot réduit, donc ils sont équivalents d'après ce qu'on a démontré plus haut.

Ainsi, l'équivalence formelle de deux mots du calcul considéré acquiert un sens géométrique concret, et le fait même de savoir si deux mots sont équivalents revient à la résolution d'un problème géométrique. De plus, l'algorithme

que nous avons décrit apparaît comme une méthode générale de résolution de tout problème géométrique d'un type donné.

Il en va de même pour les autres calculs, où l'équivalence formelle admet aussi une interprétation concrète géométrique, algébrique ou autre. On peut dire sans exagérer que chaque branche des mathématiques possède des théorèmes qui, une fois modifiés, peuvent être énoncés comme une assertion relative à l'équivalence de deux mots dans un certain calcul. Faute de place nous ne nous attarderons pas sur ces problèmes ; certaines explications supplémentaires seront données dans la suite de l'exposé (voir § 6).

Remarquons encore que l'interprétation géométrique que nous avons donnée du problème des mots dans le calcul considéré permet maintenant de construire immédiatement un algorithme et même d'une façon un peu plus simple. Pour cela il suffit de réaliser pour chacun des deux produits donnés la suite des automorphismes correspondants (ne serait-ce que sur un dessin, et de comparer les résultats obtenus.

● EXERCICE. Résoudre le problème des mots pour le calcul associatif donné dans l'alphabet $\{a, b\}$ avec les substitutions permises :

$$aaa = bb$$

$$bbbb = \Lambda.$$

§ 4. MACHINE À CALCULER À COMMANDE AUTOMATIQUE

La construction d'un algorithme pour des problèmes d'un type donné (et en particulier, celle d'un « bon » algorithme) — quand cela est possible — est généralement liée à des raisonnements subtils et compliqués, qui demandent une haute qualification et beaucoup d'ingéniosité. Mais

dès le moment où un tel algorithme est construit, le processus de résolution peut être réalisé par une personne qui n'aurait même aucune idée sur la nature du problème considéré. Il lui faut seulement savoir effectuer les opérations élémentaires et peu nombreuses dont se compose le processus et se laisser en outre guider, consciencieusement et docilement, par l'instruction (algorithme) proposée. En opérant, comme on dit parfois, d'une façon purement machinale, cette personne résoudrait avec le même succès n'importe quel problème du type considéré. L'expression « acte machinal » n'est employée ici que pour expliquer le caractère déterminant d'un algorithme ; cependant, elle prend son sens premier dans le développement actuel de la science et de la technique. En effet, à la place d'un homme hypothétique qui résout un problème dont il ne comprend même pas le sens (ou ne veut pas le connaître), on peut réellement installer une machine réalisant le même processus. Ce sera le cas pour la machine à calculer moderne à commande automatique.

Nous allons examiner maintenant les principes fondamentaux de construction et de fonctionnement de telles machines. Revenons au préalable à l'étude du processus algorithmique réalisé par l'homme (calculateur).

Guidé par l'algorithme, le calculateur réalise un processus au cours duquel s'effectuent la réception, le stockage, le traitement et la remise de certains renseignements (*information*). Habituellement, le calculateur inscrit (représente) ces renseignements sur un papier, à l'aide de chiffres, de lettres ou d'autres symboles. L'ensemble de tous ces symboles est appelé *alphabet*. Ainsi, par exemple, on se sert en algèbre d'un alphabet qui comprend, en plus de lettres usuelles, des chiffres, des symboles d'opérations algébriques, des parenthèses, etc.

Le processus calculatoire auquel participe l'homme-calculateur est caractérisé par les trois facteurs suivants (voir le schéma sur la fig. 4,*a*) :

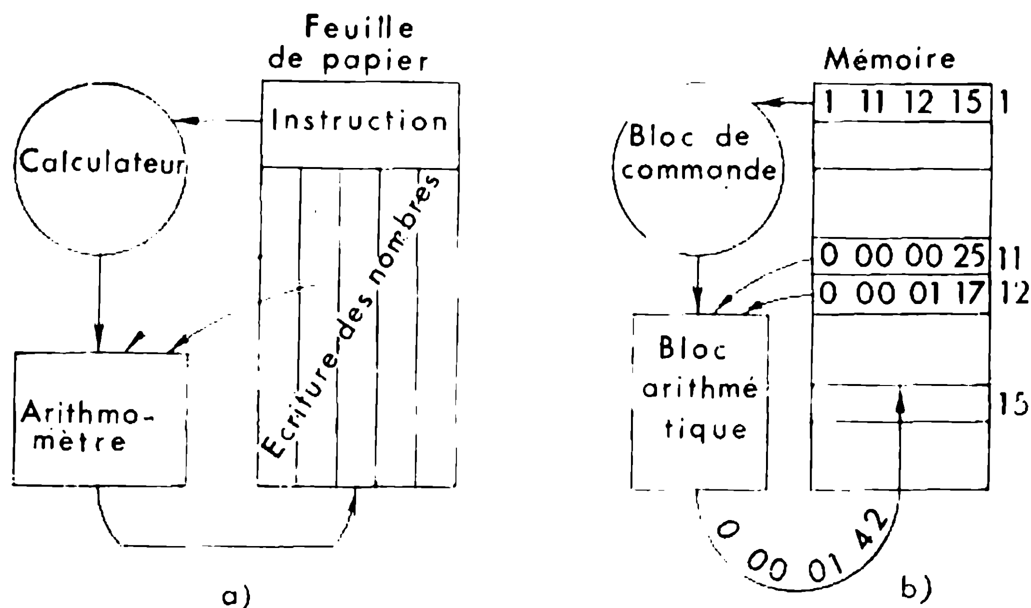


Fig. 4

1. Le *stockage de l'information* est assuré par la mise sur une feuille de papier de tous les renseignements y compris l'instruction (schéma de l'algorithme) pour la résolution du problème. Il est à noter qu'en réalité le calculateur n'enregistre pas tout sur le papier ; il se souvient de certaines choses (il les garde en mémoire, et non sur le papier) et prend certains renseignements dans différentes tables et aide-mémoires. Néanmoins, l'essentiel est que le processus calculatoire suppose l'existence de moyens qui assurent le stockage des renseignements nécessaires. Ainsi, par feuille de papier il faut entendre l'ensemble de tous les moyens qui assurent le stockage des renseignements.

2. Le *traitement de l'information* suppose que le calculateur soit capable d'effectuer les opérations élémentaires isolées prévues par l'algorithme et qu'il puisse les faire à l'aide de mécanismes spécialisés (ainsi, par exemple, les opérations arithmétiques sur les nombres peuvent être effectuées sur un arithmomètre). Chaque opération prise isolément consiste

à puiser certains renseignements (par exemple, des nombres), et conformément à l'instruction, dans des colonnes déterminées de la feuille de papier ; ceux-ci passent ensuite sur un appareil (arithmomètre) et le résultat s'inscrit en un endroit bien déterminé de la feuille.

3. La *commande du processus*, c'est-à-dire la prise d'une décision quant à la réalisation, à une étape donnée du processus, d'une opération ou d'une autre, ainsi que l'établissement des conditions dans lesquelles cette opération sera effectuée, incombe au calculateur lui-même, conformément à l'instruction.

De quels organes se compose une machine, et comment ceux-ci agissent-ils les uns sur les autres? Le fait que dans cette machine doivent se développer des processus identiques à ceux que nous venons de décrire, mais cette fois sans l'intervention du calculateur, suggère la réponse à cette question.

Premièrement, pour symboliser l'information, la machine doit disposer d'un certain alphabet ; mais à la place de la représentation graphique habituelle par des symboles que seule leur écriture distingue, les différents symboles de l'alphabet sont représentés dans la machine par certains états physiques distincts, tels que des intensités électriques ou des degrés d'aimantation différents.

Une série de considérations justifie l'avantage d'utiliser un alphabet composé de deux symboles (alphabet *binaire*) appelés de façon conventionnelle 0 et 1. Ainsi, par exemple, un tel alphabet peut être réalisé très simplement sous la forme de deux circuits: haute tension (ou passage du courant) et basse tension (le courant ne passe pas). Il faut tenir également compte de ce que les opérations logiques les plus simples sont à effectuer sur des variables qui ne peuvent prendre que deux valeurs: « vrai » ou « faux ». Néanmoins le choix d'un alphabet ou d'un autre, ainsi que le procédé de représentation dans cet alphabet des renseignements nécessaires sont loin d'être décisifs pour comprendre le mécanisme et le

fonctionnement de la machine. C'est pourquoi, sans entrer dans les détails, nous nous bornons à noter que dans les machines modernes on utilise essentiellement des alphabets et systèmes numériques binaires (au lieu du système décimal).

Ainsi, l'information qui entre dans la machine et celle qui en ressort dans son processus de fonctionnement sont représentées sous la forme de certains paramètres physiques. Dans les cas qui nous intéressent tous les renseignements dont se compose l'information seront représentés par des nombres. En particulier, l'algorithme lui-même qui doit « guider » la machine dans son fonctionnement sera également donné sous une telle forme. Les algorithmes créés spécialement pour les machines s'appellent d'habitude *programmes*. Un programme est la partie fondamentale de l'information qu'utilise la machine.

En conformité avec le schéma de la figure 4,*a* chaque machine possède des organes dont la fonction est de stocker et de traiter l'information, de commander le processus, etc. (voir le schéma de la fig. 4,*b*).

1. La *mémoire* joue le rôle d'une feuille de papier sur laquelle on fixe, dans le langage convenu de la machine, tous les renseignements nécessaires, y compris le programme. La possibilité de réaliser physiquement un mécanisme appelé à remplir ces fonctions, ne fait sûrement aucun doute. En effet, une bande magnétique, sur laquelle on fixe les renseignements codifiés pour les extraire ensuite, à peu près comme sur un magnétophone ordinaire, peut remplir ces fonctions. La *mémoire de la machine* se compose d'un ensemble de cases numérotées par les entiers naturels: 1, 2, 3, ... Ces nombres s'appellent *adresses* des cases. Chaque case peut stocker ou recevoir une information codifiée. Dans les machines à calculer chacune de ces informations est représentée, comme nous l'avons déjà dit, sous la forme d'un nombre.

Dans les machines électroniques, outre la bande magnétique, on utilise aussi dans la pratique d'autres moyens de « mémorisation », à savoir des tubes à rayons électroniques dont le principe de fonctionnement rappelle en quelque sorte celui des tubes de télévision, et le tambour magnétique; les fonctions de mémorisation sont le plus rationnellement réparties entre ces différents organes. Néanmoins, nous n'allons pas distinguer dans la suite les différentes parties composant une mémoire, et, sans nuire à la compréhension de l'exposé, nous partirons de la représentation d'un mécanisme unique de mémoire, ne serait-ce que sous la forme d'une bande magnétique.

2. Le *bloc arithmétique* joue le même rôle qu'un arithmomètre usuel bien que les principes physiques sur lesquels repose sa construction en soient bien différents. Dans le bloc arithmétique s'effectuent toutes les opérations nécessaires à la résolution d'un problème (addition des nombres, par exemple). Cela se passe par transformation dans un mécanisme électronique des signaux électriques d'entrée (correspondant aux données introduites) en signaux électriques correspondant aux données de sortie. Les signaux d'entrée, stockés dans les cases de la mémoire, gagnent le bloc arithmétique, et le signal de sortie sort du bloc pour être dirigé dans la case où il sera mis en réserve. Ce processus est schématisé sur la figure 4,*c* où les nombres des cases 11 et 12 sont additionnés et où le résultat final est adressé dans la case 15. Pour que cette opération soit effectuée au cours d'une seule passe il faut qu'avant sa réalisation les cases 11 et 12 soient reliées au bloc arithmétique et que celui-ci soit relié à la case 15; il faut aussi que l'arithmomètre soit branché sur l'opération nécessaire (dans le cas considéré sur l'addition). Mais cela relève déjà de la « compétence » du bloc de commande.

3. Le *bloc de commande* est destiné à remplir les fonctions que remplit le calculateur lui-même sur le schéma de la

figure 4,*a*. Plus précisément, à chaque étape du travail de la machine, l'*organe de commande* prépare la réalisation de l'opération suivante. Il fonctionne tout comme une station téléphonique automatique reliant les « abonnés » (centres et cases de la machine) qui interviennent dans une opération isolée. D'une façon plus imagée, le bloc de commande consulte le programme et, en fonction de celui-ci, commande le travail commun des organes de la machine qui doivent effectuer l'opération suivante.

Soulignons, pour préciser cette idée, que chaque machine est munie d'un système de commandes (ordres) parfaitement déterminé, qu'elle peut effectuer. Tout programme introduit dans la machine représente une certaine combinaison de commandes et de certains nombres auxiliaires (paramètres) placés dans les cases de la « mémoire ». Ainsi, la machine à calculer (B.E.S.M.) de l'Académie des Sciences de l'U.R.S.S. possède un système de commande dit « à trois adresses », dont chacune représente une suite de quatre nombres :

$$\alpha\beta\gamma\delta,$$

le premier désigne le numéro de l'opération à effectuer; les deux suivants, les adresses des deux cases dont le contenu est soumis à l'opération considérée; le dernier, l'adresse de la case où sera placé le résultat final (trois adresses en tout).

Pratiquement chaque commande s'inscrit dans une case sous la forme d'un seul nombre dont les chiffres se divisent en quatre groupes qui ont chacun une désignation particulière. Ainsi, par exemple, sur la figure 4,*b* le nombre 1 11 12 15 inscrit sur la case 1 de la mémoire représente l'ordre codifié suivant :

« *Additionner (opération n° 1) les nombres des cases 11, 12 et envoyer le résultat dans la case 15.* »

(La division en groupes se fait ici de droite à gauche en groupes de deux. Pour fixer les idées, nous nous en tiend-

rons dans la suite à cette méthode d'écriture des commandes.)

Un tel système se compose généralement de plusieurs dizaines de commandes. Nous ne mentionnerons ici que les plus fréquentes.

1. *Commandes arithmétiques*:
   a) 1 β γ δ — *additionner* le nombre de β avec celui de γ et envoyer la somme obtenue dans δ;
   b) 2 β γ δ — *soustraire* du nombre de β celui de γ et envoyer la différence dans δ;
   c) 3 β γ δ — *multiplier* le nombre de β par celui de γ et envoyer leur produit dans δ;
   d) 4 β γ δ — *diviser* le nombre de β par celui de γ et envoyer leur quotient dans δ.

2. *Transfert de commande*:
   e) 5 00 00 δ — passer à la commande mise en réserve dans δ (transfert *inconditionnel*);
   f) 5 01 γ δ — passer à la commande mise en réserve dans δ à condition que le nombre mis en réserve dans la case γ soit positif;
   g) 5 02 γ δ — passer à la commande mise en réserve dans δ à condition que le nombre mis en réserve dans la case γ soit négatif.

3. *Commande d'arrêt*:

0 00 00 00.

Outre les commandes énumérées ci-dessus, il en existe d'autres, parmi lesquelles celles dites d'*opérations logiques*, que nous n'allons pas examiner ici. Les commandes énumérées plus haut suffisent largement à composer toutes sortes de programmes; certains exemples seront donnés dans le paragraphe suivant.

Les commandes *f-g* s'appellent *conditionnelles*; elles ne peuvent être réalisées que si la condition correspondante est satisfaite, sinon l'organe de commande n'en tient pas compte et passe aux opérations suivantes.

Les ordres sont, en général, effectués dans l'ordre selon lequel ils ont été inscrits dans les cases de la mémoire; cet ordre ne peut être modifié qu'à la suite d'un transfert de commande (conditionnel ou non dans le cas où la condition correspondante est remplie).

Le travail de la machine se répartit en *temps*; au cours de chaque temps se réalise une commande. Au début de chaque temps, le nombre (commande) contenu dans une case de la mémoire passe de cette case dans le bloc de commande. Dès que la commande parvient dans le bloc, les liaisons correspondantes sont automatiquement assurées; par là même s'effectue l'opération suivante du processus. Ensuite, une nouvelle commande arrive dans le bloc, et la machine se lance dans l'opération suivante, etc., jusqu'à l'arrêt final de la machine.

La possibilité de réaliser techniquement un tel bloc de commande ne présente rien d'inattendu, car une telle construction n'exige plus que les capacités de n'importe quelle station automatique téléphonique, dans laquelle chaque numéro est composé au moyen d'un signal électrique correspondant.

Tels sont les trois organes fondamentaux d'une machine à calculer. Il en existe, à vrai dire, bien d'autres, tels que les organes qui envoient les données dans la machine, ou ceux qui en font sortir les résultats. Néanmoins ces organes ne sont pas d'une importance particulière pour examiner les principes de fonctionnement d'une machine et expliquer ses possibilités logiques et mathématiques. Nous pouvons donc faire abstraction de ces organes en supposant que la transmission de l'information à la machine et la sortie de l'information finale se font directement dans la mémoire.

§ 5. PROGRAMME (ALGORITHME POUR MACHINE)

Nous considérons ici des exemples de programmes établis pour une machine à calculer électronique dont le système de commandes comporte trois adresses. Ces programmes s'obtiennent à partir des algorithmes examinés plus haut; on les transforme en tenant compte de ce qu'ils ne seront plus maintenant réalisés par un homme, mais par une machine automatique munie du système de commandes considéré. De tels exemples permettent d'élucider le sens réel de l'expression « *le programme introduit détermine le travail de la machine* ». Plus précisément, nous allons montrer comment est réglé l'ordre d'entrée des commandes dans le bloc et comment on arrive, à l'aide d'un nombre relativement restreint de commandes, à effectuer parfois une longue chaîne d'opérations nécessaires pour résoudre l'une ou l'autre variante du problème envisagé.

Dans ce qui suit, on suppose que les opérations arithmétiques d'addition, de soustraction, de multiplication et de division sont désignées par les chiffres 1, 2, 3, 4 respectivement (voir § 4).

● EXEMPLE 1. Composons le programme de résolution du système d'équations

$$ax + by = c,$$
$$dx + ey = f.$$

Pour fixer les idées, convenons de placer les coefficients a, b, c, d, e, f, dans cet ordre, dans les cases de la mémoire, et à partir du numéro 51 :

Adresse	Contenu
51	a
52	b
53	c
54	d
55	e
56	f

Les cases 31-50 sont réservées pour des résultats intermédiaires et définitifs.

Comme on le voit sur les formules

$$x = \frac{ce - fb}{ae - bd}; \quad y = \frac{af - cd}{ae - bd},$$

il faut, pour obtenir le résultat final, effectuer successivement 6 multiplications, 3 soustractions et 2 divisions. Le programme que nous proposons ici se compose donc de 12 commandes inscrites dans les cases 1-12:

Adresse	Contenu (commandes)
1	3 53 55 31
2	3 56 52 32
3	3 51 56 33
4	3 53 54 34
5	3 51 55 35
6	3 52 54 36
7	2 31 32 37
8	2 33 34 38
9	2 35 36 39
10	4 37 39 40
11	4 38 39 41
12	0 00 00 00

Les commandes sont introduites dans le bloc de commande et exécutées dans l'ordre croissant de leurs adresses. La dernière commande une fois réalisée, on stocke dans les cases 31-41 les nombres suivants:

Adresse	Contenu
31	ce
32	fb
33	af
34	cd
35	ae
36	bd
37	$ce - fb$
38	$af - cd$
39	$ae - bd$
40	$\frac{ce - fb}{ae - bd}$
41	$\frac{af - cd}{ae - bd}$

qui représentent les résultats intermédiaires des calculs et le résultat final (dans les cases 40, 41).

● EXEMPLE 2. Supposons qu'on doive trouver les solutions de n systèmes donnés d'équations:

$$\begin{cases} a_i x + b_i y = c_i \\ d_i x + e_i y = f_i \end{cases} \quad i = 1,\ 2,\ \ldots,\ n.$$

L'algorithme de résolution apparaît comme la répétition de n algorithmes de résolution d'un seul système de ce type, et on pourrait facilement réaliser le programme correspondant pour la machine de la manière indiquée plus haut, en sorte que les $6\,n$ coefficients sont répartis dans les $6n$ cases de la mémoire et que le programme soit composé de $11n + 1$ commandes. Après que le premier cycle de 11 commandes ait donné la solution du premier système, le second cycle de 11 commandes donne celle du deuxième, etc., n fois; la commande $(11n + 1)$ est la commande d'arrêt.

Il n'est pas rationnel d'augmenter de la sorte le volume du programme, et l'on peut s'en dispenser. Pour cela, remarquons que chaque cycle composé de 11 commandes peut être obtenu du cycle qui le précède par un changement des adresses qui interviennent dans les commandes. En effet, si les $6n$ coefficients occupent chacun une case, à la suite les uns des autres, en commençant, par exemple, au numéro 51, alors, pour obtenir les commandes correspondantes du deuxième cycle, il faut augmenter de 6 les adresses des facteurs dans les six premières commandes.

Ensuite, pour que la solution du système suivant soit inscrite dans les deux cases qui suivent immédiatement celles qui conservent la solution du système précédent, il faut augmenter de deux chacune des dernières adresses des commandes 10 et 11. Une telle modification des adresses

s'avère possible grâce aux commandes dites de *réadressage* dont le nombre (dans le cas considéré) est de 8. Mettons pour cela dans les cases 25 et 27 les paramètres:

Numéro de la case	Contenu
25	0 06 06 00
26	0 00 00 02

et dans les cases 12-18 les huit commandes de réadressage:

Numéro de la case	Contenu
12	1 01 25 01
13	1 02 25 02
14	1 02 25 03
15	1 04 25 04
16	1 05 25 05
17	1 06 25 06
18	1 10 26 10
19	1 11 26 11

Les commandes des cases 1-19 étant effectuées, les cases 40-41 seront prises, comme auparavant, pour stocker la solution du premier système d'équations, mais dans les cases 1-6 et 10-11 seront maintenant placées les commandes réadressées suivantes:

Numéro de la case	Contenu
1	3 59 61 31
2	3 62 58 32
3	3 57 62 33
4	3 59 60 34
5	3 57 61 35
6	3 58 60 36
10	4 37 39 42
11	4 38 39 43

Ainsi, si l'on effectue maintenant une fois de plus les commandes des cases 1-19, on obtiendra, après le deuxième

cycle de fonctionnement de la machine, la solution du deuxième système d'équations qui sera stockée dans les cases 42-43; en outre, les commandes 1-6 et 10-11 seront réadressées pour permettre un troisième cycle analogue du fonctionnement, etc.

La question se pose donc de savoir comment on peut assurer la réalisation cyclique des 19 commandes autant de fois qu'il y a de systèmes d'équations, en sorte que, toutefois, la machine s'arrête une fois qu'on a toutes les solutions. Mettons pour cela dans les cases 27 et 28 les paramètres 0 00 00 01 et n (n étant le nombre de systèmes donnés) et ajoutons aux 19 commandes ci-dessus les trois commandes suivantes:

Numéro de la case	Contenu
20	2 28 27 28
21	5 01 28 01
22	0 00 00 00

La commande 20 fait diminuer de l'unité le contenu de la case 28 après la réalisation du cycle des commandes 1-19.

La commande 21 représente le transfert conditionnel à la commande 1: celui-ci se fait tant que la case 28 conserve un nombre positif, ce qui assure le transfert au cycle suivant des commandes 1-19 appelées à résoudre le système d'équations suivant. Si dans la case 28 figure 0, ce qui arrive exactement après n cycles 1-19, le transfert de commande ne se fait plus, il est alors remplacé par la commande 22 d'arrêt de la machine.

Ainsi, on comprend que le programme constitué des commandes 1-22 indiquées plus haut, compte tenu des paramètres introduits dans les cases 25-28, soit adapté à la résolution du problème posé de n systèmes d'équations. Il est commode de représenter la structure d'un tel programme sous la forme du schéma suivant:

	Commandes
1 2 3 4 5 6 7 8 9 10 11	*Opérations arithmétiques*
12 13 14 15 16 17 18 19	*Opérations de réadressage*
20	*Compte des cycles*
21	*Conditions de transfert des commandes*
22	*Arrêt*

	Paramètres			
25	0	06	06	00
26	0	00	00	02
27	0	00	00	01
28		n		

● EXEMPLE 3. Considérons maintenant le programme de recherche du P.G.C.D. de deux nombres a, b. Convenons de réserver les cases 12 et 13 aux données initiales a, b respectivement, et les cases 14 et 15 aux calculs intermédiaires, pour stocker ensuite le résultat final après l'arrêt de la machine dans la case 15. Le programme proposé ci-dessous est conforme à la formulation de l'algorithme d'Euclide donnée au § 1.

Numéro de la case	Contenu (commande)	Explications
01	1 12 05 15	Transfert du nombre de 12 dans 15
02	2 12 13 14	Différence des nombres de 12 et 13 est envoyée dans 14
03	5 02 14 06	Passage à la commande 06, si dans 14 est stocké un nombre négatif
04	5 01 14 09	Passage à la commande 09, si dans 14 est stocké un nombre positif
05	0 00 00 00	Arrêt
06	1 13 05 12	Transfert du nombre de 13 dans 12
07	1 15 05 13	Transfert du nombre de 15 dans 13
08	5 00 00 01	Passage inconditionnel à la commande 01
09	1 13 05 12	Transfert du nombre de 13 dans 12
10	1 14 05 13	Transfert du nombre de 14 dans 13
11	5 00 00 01	Passage inconditionnel à la commande 01

Après les deux premiers temps, dans les cases 12-15 sont stockés les nombres suivants:

Numéro de la case	Contenu
12	a
13	b
14	$a - b$
15	a

Si $a - b = 0$ (c'est-à-dire $a = b$), les commandes 03 et 04 de transfert conditionnel de commande sont supprimées et on passe à la commande 05 qui arrête la machine. A ce moment-là, dans la case 15 est effectivement stocké le résultat cherché (comparer à l'indication 3 du § 1).

Si $a - b < 0$ (c'est-à-dire $a < b$), la commande 03 transmet la commande à 06 qui intervertit, avec la commande 07 qui la suit, l'ordre des nombres a, b dans les cases 12 et 13 (comparer à l'indication 4). La commande 08 assure ensuite le passage inconditionnel à 01 ; c'est par là que commence le deuxième cycle de fonctionnement de la machine.

Si $a - b > 0$, (c'est-à-dire $a > b$) la commande 03 est supprimée et 04 transfère la commande à 09 qui, avec la commande 10 qui la suit, met dans les cases 12 et 13 le nombre précédemment soustrait et le reste, c'est-à-dire b et $a - b$ respectivement (cf. avec l'indication 4). La commande 11 assure le passage inconditionnel à la commande 01 et ceci marque le commencement du deuxième cycle du fonctionnement de la machine.

La suite de cycles du fonctionnement de la machine introduit dans les cases 12 et 13 la suite de couples des nombres suivants :

$$(a_1, b_1), \ (a_2, b_2), \ \ldots, (a_i, b_i), \ (a_{i+1}, b_{i+1}), \ \ldots$$

et dans la case 15 la suite de nombres

$$a_1, \ a_2, \ a_3, \ \ldots, a_i, \ a_{i+1}, \ \ldots$$

jusqu'à ce qu'apparaisse le premier couple de nombres égaux a_k, b_k. La commande 05 arrête alors la machine et le résultat cherché a_k est placé dans la case 15.

Les exemples ci-dessus reflètent déjà bien les deux principes fondamentaux suivants du fonctionnement d'une machine à calculer automatique :

1. Les commandes d'un programme sont habituellement réalisées dans l'ordre fixé dans lequel elles sont inscrites dans les cases de la mémoire.

Néanmoins la machine peut modifier automatiquement la marche des calculs en fonction des résultats intermédiaires obtenus dans les calculs. Cela se fait à l'aide des commandes conditionnelles.

2. Avec un programme relativement simple la machine peut effectuer une chaîne assez longue de calculs grâce à sa faculté de modifier elle-même et de répéter plusieurs fois tout le programme ou certaines de ses parties. Une telle possibilité s'explique par le fait que le programme, codifié par des nombres, est stocké dans la mémoire même où figurent les nombres usuels et, par conséquent, la machine peut donc opérer aussi sur les nombres conventionnels représentant les codes des commandes. Ainsi, par exemple, la machine peut réaliser la réexpédition de certaines commandes.

Ces principes caractérisent le fonctionnement de la machine même pour la résolution de problèmes dont le caractère est loin d'être arithmétique. Ainsi, par exemple, on peut programmer l'algorithme de Thésée (recherches dans le labyrinthe) ou les algorithmes bien connus du problème des mots dans certains calculs associatifs et réaliser les processus correspondants dans la machine. Pour cela, la machine doit bien sûr savoir effectuer certaines opérations élémentaires supplémentaires ainsi que le transfert de commandes qui interviennent dans les problèmes arithmétiques considérés plus haut. On peut effectuer ces quelques types d'opérations simples dans les machines à calculer électroniques (on prévoit les commandes correspondantes). Il suffit donc de modifier le programme pour que la machine elle-même réalise l'opération voulue.

Néanmoins, non seulement les mathématiques mais également d'autres domaines fort divers de l'activité humaine sont riches en processus qui résultent d'une instruction strictement définie (i.e. de l'algorithme) et qu'on peut aussi programmer. Ainsi, par exemple, pour la comptabilité et la planification, l'examen des données, leur traitement et la

mise au point de balances matérielles pour obtenir les résultats optimaux s'effectuent à l'aide d'une longue chaîne d'opérations élémentaires de plusieurs types dont la réalisation est strictement réglementée par des instructions et schémas spéciaux. Dans les autres cas, bien qu'un algorithme précis n'ait pas encore été mis au point, on pourrait quand même en construire un et le porter à la perfection formelle. Ceci se rattache, en particulier, au problème de traduction d'une langue dans une autre. Avec un traitement formel suffisant et une classification des règles grammaticales et stylistiques fondamentales et aussi avec des procédés permettant de se servir du dictionnaire on peut créer un algorithme tout à fait convenable pour traduire, par exemple, un texte scientifique ou d'affaire (pour certaines langues de tels algorithmes sont déjà composés).

Dans cet ordre d'idées il est encore intéressant de rappeler que dans de nombreux jeux, les règles du jeu portent les traits d'une instruction formelle à l'action ; cela permet de poser la question de l'élaboration des algorithmes pour que le jeu soit mené avec succès. Un tel algorithme, fondé sur une certaine tactique du jeu, doit aussi montrer à un joueur l'unique coup (le meilleur pour cette tactique) à jouer dans chaque position. Ainsi, par exemple, on peut munir le jeu d'échecs d'un système d'évaluation des figures ; un nombre de points très grand pour le roi, un nombre plus petit pour la reine, encore plus petit pour la tour, etc., le pion ayant le moins de valeur. On peut de plus évaluer d'une certaine façon les avantages stratégiques des figures (la disposition des figures sur l'échiquier situées plus près de son centre, leur mobilité, etc.). La différence entre la somme des points des figures blanches et celle des figures noires caractérise pour une tactique donnée la supériorité matérielle et stratégique des blanches sur les noires à un stade quelconque de la partie. L'algorithme le plus simple du jeu consiste donc à trier toutes les marches possibles dans une position donnée et à en choisir ensuite la plus avanta-

geuse du point de vue du système d'évaluation établi. Un algorithme plus avantageux mais plus compliqué sera celui qui consiste à passer en revue toutes les combinaisons de trois ou, disons, de cinq déplacements très courts et à choisir à partir de là la marche optimale*).

De tout ce qu'on a dit plus haut il devient clair que les aspects du travail intellectuel qui sont accomplis, ou qui peuvent l'être, en fonction de certains algorithmes, sont fort divers. Dans de tels cas on peut en principe programmer ces algorithmes et confier l'exécution du travail correspondant à une machine à commande automatique. En particulier, on a déjà élaboré les programmes de traduction d'une langue dans une autre et du jeu d'échecs; ces programmes sont réalisés avec succès par des machines, telles que la machine à calculer (B.E.S.M.) de l'Académie des Sciences de l'U.R.S.S.

§ 6. NÉCESSITÉ DE PRÉCISER LA NOTION D'ALGORITHME

Tout ce qu'on a dit plus haut révèle le lien profond entre les algorithmes et les machines à calculer automatiques.

Il est clair que chaque processus dont les étapes isolées sont successivement réalisées par une machine à calculer

---

*) De plus, il existe, du point de vue théorique, un algorithme optimal assurant toujours le gain, quand c'est possible. Pour les problèmes ayant trait au jeu d'échecs du type *mat aux noirs en N coups* l'algorithme impose aux blancs la démarche suivante: examiner toutes les séries possibles de N coups B_1, N_2, B_3, N_4, . . ., N_{n-1}, B_n, où B_1, B_3, . . ., N_2, N_4, . . ., sont les coups des blancs et des noirs respectivement et en choisir ensuite une suite de coups B_1, B_3, . . ., qui garantisse le gain indépendamment des coups N_2, N_4, . . . Néanmoins un tel algorithme est tellement encombrant que même les machines à calculer rapides modernes s'avèrent impuissantes à le réaliser.

peut être décrit par un algorithme. D'autre part, tous les algorithmes connus jusqu'à maintenant, ainsi que tous ceux qu'on peut prévoir dans l'état actuel de la science, sont *en principe* réalisables dans les machines automatiques.

Cette dernière affirmation est à préciser. Comme nous l'avons déjà dit, le processus d'application d'un algorithme à la résolution de certains problèmes d'un type donné peut s'avérer aussi long que l'on veut, et les inscriptions des renseignements sur lesquels opère l'algorithme peuvent être trop volumineuses. D'autre part, la mémoire des machines à calculer modernes est limitée en volume (car le nombre des cases est fini et la capacité de chaque case est restreinte). C'est pourquoi, dans certaines conditions, un algorithme peut s'avérer pratiquement irréalisable.

Pour illustrer cette affirmation, considérons l'algorithme d'Euclide. Le problème le plus simple consistant à rechercher le P.G.C.D. de deux nombres quelconques peut s'avérer irréalisable pour un calculateur qui le résout « à la main », si la résolution de ce problème nécessite plus de papier et d'encre que la quantité dont dispose le calculateur. D'une façon analogue, l'algorithme d'Euclide appliqué à un problème concret donné peut s'avérer irréalisable pour une machine, si le volume de la mémoire nécessaire à sa réalisation est supérieur à celui dont dispose la machine.

Comme nous l'avons souligné, dans de tels cas le processus d'application d'un algorithme est considéré comme un *processus potentiellement réalisable*, qui conduit au résultat cherché après un nombre fini (même très grand) d'étapes.

Lorsqu'on parle de la possibilité de réaliser un algorithme en machine, on a en vue la possibilité de rendre le volume de sa mémoire infiniment grand.

Le lien mentionné entre la notion d'algorithme et celle de machine automatique à mémoire potentiellement illimitée permet de mieux comprendre l'essence de chacune de ces notions. Toute précision concernant l'une de ces notions est aussi une précision pour l'autre.

Néanmoins, tout en soulignant les traits communs de ces notions, nous n'avons pas donné jusqu'à présent leurs définitions exactes. La définition mathématique rigoureuse de la notion d'algorithme (ainsi que la notion voisine de machine à calculer automatique) n'a été mise au point en science que dans les années 30. Pourquoi donc pendant si longtemps les mathématiciens ont-ils pu se tirer d'affaire avec une notion vague d'algorithme? Par quoi s'explique le besoin urgent et relativement récent d'élaborer une définition mathématiquement rigoureuse de cette notion qui pourrait devenir un objet de recherche mathématique?

Tout récemment encore, le terme d'algorithme se rencontrait en mathématiques lorsqu'il s'agissait de construire des algorithmes concrets, lorsque *l'affirmation de l'existence d'un algorithme pour la résolution des problèmes d'un type donné s'accompagnait de sa description réelle.* Dans de telles circonstances, la personne à laquelle était communiqué le système de règles formelles n'avait plus qu'à les appliquer et à se persuader qu'elles mèneraient automatiquement au résultat cherché. C'est pourquoi la question de définir rigoureusement la notion d'« algorithme » ne se posait pas, et l'on pouvait se contenter d'une représentation vague, quoique suffisamment compréhensible pour tout mathématicien. Néanmoins, l'évolution même des mathématiques a accumulé des faits qui ont littéralement bouleversé cette situation. La cause en est le désir naturel des mathématiciens de créer des algorithmes toujours plus puissants, capables de résoudre des classes plus vastes de problèmes (problèmes d'un type très général). Venons-en maintenant à l'examen de ces faits.

Rappelons-nous l'algorithme d'extraction de racine carrée qui figure dans tous les manuels scolaires. On peut formuler un problème plus général: construire un algorithme général d'extraction de racine de tout ordre d'un nombre quelconque. Il est naturel de s'attendre à ce qu'un tel algorithme soit plus difficile à construire, mais la perspective

d'en disposer est intéressante. On peut même aller plus loin. Extraire la racine d'ordre n d'un nombre a équivaut à résoudre l'équation $x^n - a = 0$ (trouver la racine de cette équation). On peut donc formuler un problème encore plus général:

Construire un algorithme permettant de trouver, pour toute équation du type

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0 \qquad (*)$$

(n étant un nombre entier arbitraire), *toutes ses racines* *).

Il va de soi qu'un tel algorithme s'avère encore plus difficile à réaliser; au fond, le contenu fondamental d'un secteur de l'algèbre supérieure, à savoir l'algèbre des polynômes, se ramène à construire et à justifier cet algorithme; mais de plus cet algorithme a une importance bien plus grande.

Les exemples étudiés caractérisent déjà, dans une certaine mesure, le désir des mathématiciens de créer des algorithmes toujours plus puissants, capables de résoudre des classes de plus en plus vastes de problèmes (problèmes de type très général). Il va de soi que, dans cet ordre d'idées, le problème de résolution d'une équation du type (*) ne représente pas encore une limite qu'on ne puisse pas franchir. De plus, si nous voulons être conséquents dans notre désir d'élargir la classe des problèmes pour lesquels il est souhaitable d'avoir un seul algorithme de résolution, nous devons obligatoirement poser le problème suivant:

Construire un algorithme permettant de résoudre tout problème mathématique.

Une telle formulation du problème est déjà si générale qu'elle peut être considérée comme un défi aux mathématiques; de plus, une telle formulation a des points faibles, ne serait-ce que parce que l'expression « tout problème mathé-

---

*) Plus précisément, trouver pour tout entier k les approximations décimales des racines par défaut et par excès à $\frac{1}{10^k}$ près.

matique » n'est pas très claire. Mais l'attraction exercée par un tel problème n'a pas besoin de justification.

Ce problème a une histoire. Le grand mathématicien et philosophe allemand Leibniz (1646-1716) rêvait déjà d'inventer une méthode universelle qui permettrait de résoudre effectivement tout problème mathématique. Bien qu'il n'ait pas réussi à créer un tel algorithme universel, Leibniz pensait quand même qu'un moment viendrait où on le trouverait et où toute discussion entre mathématiciens serait automatiquement résolue en fonction de cet algorithme universel, avec un crayon et du papier.

Par la suite ce problème fut précisé sous forme d'un des problèmes fondamentaux de la logique mathématique dit *problème de discernement de déductibilité*. Sans pouvoir donner ici un exposé détaillé du problème, bornons-nous à faire quelques remarques d'ordre général.

Comme on le sait, la méthode axiomatique en mathématiques consiste en ce que toutes les propositions (théorèmes) d'une théorie donnée s'obtiennent par déduction formelle logique à partir d'un certain nombre de propositions (axiomes) admises sans démonstration. La géométrie fut la première à se voir munie d'une base axiomatique mais aujourd'hui presque toutes les théories mathématiques modernes sont déjà axiomatisées. En logique mathématique est décrit un langage spécial des formules permettant de mettre toute proposition d'une théorie mathématique sous forme d'une formule parfaitement déterminée.

En utilisant la terminologie introduite plus haut lors de l'étude des calculs associatifs, on peut dire que toute formule de ce type représente un mot dans un alphabet spécial, qui contient, outre les symboles utilisés en mathématiques (lettres, parenthèses, etc.), d'autres symboles spécifiques désignant des opérations logiques telles que la négation, la somme logique, etc. Mais l'analogie fondamentale avec les calculs associatifs consiste en ce que le processus même de déduction logique à partir d'une prémisse R de la consé-

quence S peut être mis sous la forme de transformations formelles des mots, fort semblables aux substitutions permises dans le calcul associatif. Cela permet de parler d'un calcul logique dans lequel est indiqué le système des transformations permises qui représentent les actes élémentaires de la déduction logique à partir desquels on forme toute déduction formelle logique aussi complexe que l'on veut. Un exemple d'une telle transformation permise est la suppression de deux signes de négation voisins dans une formule ; ainsi, par exemple, le mot « non non joli» peut être transformé en celui de « joli » (il est intéressant de comparer cette transformation permise avec la substitution $aa - \wedge$ dans le calcul associatif de l'exemple 4).

Ainsi, le problème de la déduction logique de la proposition S à partir de la prémisse R dans le calcul logique considéré se ramène au problème de l'existence d'une chaîne déductive allant d'un mot qui représente la prémisse R au mot qui symbolise la proposition S. Le problème de discernement de déductibilité peut alors être formulé comme suit :

pour deux mots quelconques (formules) R et S du calcul logique, existe-t-il ou non une chaîne déductive allant de R à S.

La solution est recherchée sous la forme d'un algorithme donnant la réponse à toutes les questions de ce type (R et S sont arbitraires).

Il est aisé de comprendre que la construction d'un tel algorithme représenterait une méthode universelle pour la résolution automatique des problèmes les plus variés de toutes les théories mathématiques qui sont construites axiomatiquement. En effet, la justesse d'une affirmation S quelconque (la formulation d'un théorème quelconque, par exemple) signifie dans une telle théorie la possibilité de sa déduction logique à partir d'un système d'axiomes pris pour la prémisse R. Mais, en faisant alors appel à l'algorithme de la déductibilité, on pourrait établir si l'affirmation S est vraie ou non dans la théorie considérée. De plus,

dans le cas d'une réponse affirmative on pourrait déterminer effectivement, dans le calcul logique, la chaîne déductive correspondante, et établir en fonction de celle-ci la chaîne des déductions que composent la démonstration de l'affirmation considérée. Un tel algorithme permettrait au fond de résoudre par une méthode effective unique, presque tous les problèmes mathématiques posés et non encore résolus jusqu'à présent. Ce fait rend donc compréhensibles non seulement l'intérêt de construire un tel « algorithme universel » et, par suite, une machine « toute-puissante », mais également les difficultés de cette construction.

En effet, malgré les efforts persévérants et continus de grands spécialistes, on n'a pu encore surmonter les difficultés liées à une telle construction. De plus, même lorsqu'on a tenté de construire un algorithme pour la résolution de certains problèmes mathématiques d'un type plus particulier, on a rencontré des difficultés semblables ; c'est le cas du problème d'Hilbert sur les équations diophantiennes (voir § 1) et d'une série d'autres problèmes que nous examinerons plus loin.

Les tentatives, nombreuses mais vaines, pour créer ces algorithmes ont permis de comprendre que ces difficultés se situent au niveau des principes, et ont conduit à la conclusion selon laquelle un *algorithme de résolution ne peut pas être construit pour toute classe de problèmes.*

Il est clair que l'affirmation sur l'insolubilité algorithmique d'une certaine classe de problèmes, c'est-à-dire sur l'impossibilité de donner l'algorithme correspondant, n'est pas la simple constatation du fait qu'un tel algorithme ne nous est pas connu et que personne ne le trouvera. Une telle affirmation est en même temps une prédiction selon laquelle un tel algorithme ne sera jamais trouvé (en d'autres termes, qu'il n'existe pas) ; elle a besoin d'être justifiée par une démonstration mathématique. Toutefois, parler d'une telle démonstration n'a pas de sens, tant qu'on n'a pas de définition exacte de la notion d'« algorithme », sinon on ne sait

pas très bien la non-existence de quoi on veut démontrer. Il est utile de rappeler que l'histoire des mathématiques connaît certains problèmes qui étaient restés longtemps sans solution et pour lesquels on a établi par la suite l'impossibilité d'une résolution par les moyens qu'on voulait utiliser. Tels sont, par exemple, les problèmes de trisection d'un angle et de résolution d'une équation par radicaux.

Depuis l'école on sait comment faire la bisection d'un angle à l'aide du compas et de la règle. Même les grecs de l'Antiquité ont déjà posé un problème analogue sur la trisection d'un angle à l'aide de ces instruments. Néanmoins on a démontré par la suite l'impossibilité d'une telle résolution du problème.

On sait aussi depuis l'école que les racines d'une équation du second degré s'expriment au moyen de ses coefficients à l'aide d'une formule où figurent les symboles des opérations arithmétiques et le signe du radical (du second degré). Pour les équations du troisième et quatrième degré on a également établi des formules s'exprimant par radicaux, qui sont, il est vrai, plus compliquées et comportent des radicaux « à plusieurs étages ». Néanmoins toutes les tentatives poursuivies même au début du XIX[e] siècle pour trouver des formules analogues par radicaux pour les équations de degré supérieur à quatre se sont avérées vaines jusqu'à ce qu'on ait finalement établi le résultat remarquable suivant :

Pour n supérieur ou égal à 5 il n'existe pas de formule pour exprimer par radicaux les racines d'une équation de degré n au moyen de ses coefficients.

Dans les deux cas les démonstrations de l'impossibilité de résoudre les problèmes en question se sont avérées possibles, car on avait deux définitions précises, répondant aux questions : « que veut dire construire à l'aide d'un compas et d'une règle ? » et « que signifie résoudre une équation par radicaux ? » Il est à noter que ces deux définitions précisent aussi le sens de certains algorithmes d'un *type spécial*, à savoir l'algorithme de résolution des équations *par radicaux*

(et non pas un algorithme général de résolution des équations) et l'algorithme de trisection de tout angle à l'aide du compas et de la règle (et non pas un algorithme général de trisection). Mais tout récemment encore, on n'avait pas de définition générale d'*algorithme*. C'est pourquoi le problème de mettre au point cette notion est devenu donc l'un des problèmes fondamentaux des mathématiques modernes. Il importe de souligner que la notion d'algorithme (de même que toute autre définition mathématique) ne doit pas être considérée comme un simple accord des mathématiciens sur une compréhension universelle du terme d'algorithme. Lorsqu'on formule une telle définition, il faut contourner la difficulté qui réside en ce que cette définition doit refléter d'une façon correcte l'essence même d'une notion qu'on avait déjà en fait, bien que sous une forme vague, et qu'on a illustré plus haut sur de nombreux exemples. A cette fin on a entrepris, à partir des années 30, toute une série de recherches ayant pour but de révéler tous les moyens utilisés en fait pour la construction des algorithmes. Le problème consistait donc à définir sur cette base une notion d'algorithme correcte non seulement du point de vue de sa précision formelle, mais, et c'est important, du point de vue de la correspondance effective de l'essence de la notion aussi définie. Différents chercheurs ont pris pour point de départ diverses considérations techniques et logiques, ce qui a permis d'obtenir plusieurs définitions de la notion d'*algorithme*. Néanmoins elles se sont avérées toutes équivalentes entre elles et définissent ainsi une seule et même notion ; on a donc obtenu la notion précise d'*algorithme* moderne.

Le fait même que tous les procédés, bien que fort différents, utilisés pour préciser la notion d'*algorithme* conduisent toujours au fond à un seul et même résultat est particulièrement important ; il témoigne justement de l'intérêt de la mise au point de cette notion.

La définition d'algorithme faisant voir l'essence même des processus réalisés dans une machine est, d'un intérêt

particulier du point de vue des mathématiques sur machine.

Pour obtenir une telle définition mathématiquement rigoureuse, il faut représenter le mécanisme du fonctionnement d'une machine sous la forme d'un schéma standard, d'une structure logique la plus simple possible, mais suffisamment précise, pour faire l'objet d'une étude mathématique. Cela fut fait pour la première fois par le mathématicien anglais Turing qui proposa une conception très simple, et en même temps très générale, de machine à calculer. Il est à souligner que la machine de Turing a été décrite en 1937 ; c'est-à-dire avant l'apparition des calculateurs modernes. Son créateur partait de l'idée générale consistant à comparer le fonctionnement d'une machine au travail d'un homme opérant selon une instruction rigoureuse. L'exposé des problèmes que nous donnons ci-dessous utilise déjà les principes fondamentaux du fonctionnement des machines à calculer électroniques.

§ 7. MACHINE DE TURING

Les traits distinctifs de la machine de Turing par rapport aux machines à calculer électroniques (voir §§ 4-5), sont :

1. Dans la machine de Turing, la division du processus de calcul en opérations élémentaires est poussée, dans un certain sens, à l'extrême. Ainsi, par exemple, l'opération d'addition qui figure dans une machine électronique comme la seule opération élémentaire est ici divisée en une chaîne d'opérations encore plus simples.

Il va de soi que cela rend le processus de calcul bien plus lent dans la machine de Turing mais simplifie considérablement sa structure logique et lui donne une forme standard très facile à examiner théoriquement.

2. Dans la machine de Turing, une partie de la mémoire *) est représentée sous la forme d'une bande illimitée divisée en cases.

Il est clair qu'aucune machine réelle ne peut avoir de mémoire illimitée (bande illimitée), et en ce sens la machine de Turing ne représente qu'un schéma idéalisé qui reflète la possibilité d'augmenter le volume de la mémoire.

Une telle idéalisation est justifiée si l'on prend en considération le lien mentionné plus haut entre la notion d'algorithme et celle de machine à mémoire potentiellement illimitée.

Passons maintenant à la description détaillée de la machine de Turing.

1. Elle dispose d'un nombre fini de signes (symboles)

$$s_1, s_2, \ldots, s_k$$

formant l'*alphabet* dit *extérieur* destiné à coder les données introduites dans la machine et celles qui en sortent. Pour généraliser notre exposé, convenons de considérer que parmi les symboles de l'alphabet extérieur il existe un *symbole vide* (disons s_1) dont la mise dans une case quelconque de la bande (de la mémoire) supprime le symbole qui y était auparavant stocké et la laisse vide. Nous dirons que dans une case vide est stocké le symbole vide.

Dans chaque case ne peut être stocké, à chaque étape du fonctionnement de la machine, qu'un seul signe. Tout renseignement enregistré sur la bande est représenté par un nombre fini de symboles de l'alphabet extérieur, différant du symbole vide et mis dans certaines cases de la mémoire à raison de 1 par case. Le fonctionnement de la machine commence par l'introduction de l'information initiale; sa marche peut être divisée en une série de manipulations successives appelées à transformer l'information

*) La mémoire dite extérieure.

initiale en informations intermédiaires (à la fin de chaque manipulation, l'ensemble des symboles enregistrés sur la bande constitue l'information intermédiaire). Comme information initiale introduite on peut prendre tout système fini de signes de l'alphabet extérieur (tout mot de cet alphabet) arbitrairement disposés dans les cases de la mémoire. Mais suivant la façon dont a été introduite l'information initiale $\mathfrak{A}$, les deux cas suivants peuvent se présenter:

a) après avoir effectué un nombre fini de manipulations la machine s'arrête en émettant le signal d'arrêt; dans ce cas une certaine information $\mathfrak{B}$ est représentée sur la bande. On dit alors que la machine est applicable à l'information initiale $\mathfrak{A}$ et la transforme en l'information résultante $\mathfrak{B}$;

b) l'arrêt et le signal d'arrêt n'apparaissent jamais. On dit alors que la machine n'est pas applicable à l'information initiale $\mathfrak{A}$.

On dit que la *machine résout une certaine classe de problèmes* si elle est toujours applicable à l'information qui réalise dans un certain code les conditions de tout problème de ce type pris isolément, et la transforme en l'information qui représente, dans le même code, la solution de ce problème.

2. Le système de commandes à trois adresses, propre à beaucoup de machines à calculer électroniques, résulte de l'existence d'opérations élémentaires nécessitant trois cases de la mémoire à la fois.

Mais dans certaines machines à calculer électroniques on utilise un système de commandes à une seule adresse; chaque manipulation de la machine n'occupe alors qu'une seule case de la mémoire (nous l'appellerons dans la suite *case examinée* à l'étape donnée). Ainsi, par exemple, la commande à trois adresses de l'addition des nombres des cases β, γ avec inscription du résultat dans la case δ peut être remplacée par les trois commandes consécutives suivantes: a) appel du nombre figurant dans la case β; b) appel du nombre de la case γ; c) mise du résultat dans la case δ.

Dans la machine de Turing, le système des opérations élémentaires ainsi que celui des commandes à une adresse sont encore plus simplifiés : lors de chaque manipulation prise isolément la commande est uniquement destinée à remplacer le seul signe s_i stocké dans la case visualisée par un autre signe s_j. Pour $j = i$ cela signifie que le contenu de la case considérée reste le même ; pour $j = 1$ cela signifie que si la case considérée contenait un signe celui-ci est à effacer. La simplification ultérieure consiste en ce que, lors du passage de la machine d'une manipulation à la suivante, l'adresse de la case considérée ne peut augmenter ou diminuer que d'une seule unité, c'est-à-dire qu'on examine la case voisine, à gauche ou à droite, ou bien encore la même case.

L'idée de cette simplification consiste en ce qu'on recherche le contenu d'une case isolée nécessaire au processus par une vérification successive de toutes les cases jusqu'à ce que l'on trouve la case cherchée. Cela rend le processus de calcul bien plus lent, mais on a l'avantage suivant : les adresses des cases figurant dans les commandes d'un programme peuvent être réduites aux trois adresses standards suivantes dont voici les symboles :

D-examiner la case voisine à droite,
G-examiner la case voisine à gauche,
I-examiner la même case.

3. Pour traiter l'information numérique stockée dans la mémoire, la machine électronique décrite dans les §§ 4-5 possède un bloc arithmétique $\mathcal{L}$ qui peut avoir l'un des états finis suivants : addition, soustraction, etc.

Pour effectuer dans ce bloc une opération quelconque, on y introduit par certains canaux non seulement les nombres sur lesquels s'effectuent les opérations mais également des signaux commandant l'opération correspondante (cf. fig. 4,*b*). Dans la machine de Turing le traitement de l'information s'effectue dans un *bloc logique* qui peut aussi avoir un nombre

fini d'états, soit

$$q_1, q_2, \ldots, q_m$$

les signes spéciaux désignant ces opérations. Ce bloc a deux canaux d'entrée ; l'un sert à introduire à chaque stade du fonctionnement de la machine (manipulation) un signe de la case considérée, l'autre le signe q_l de l'état assigné au bloc dans une manipulation donnée ; par le canal de sortie le bloc adresse dans la case considérée le signe correspondant « transformé » s_j qui est une fonction bijective des signaux s_j, q_l introduits à l'entrée. Les commandes qui assurent le fonctionnement de la machine à chaque manipulation isolée sont donc de la forme :

$$Dq_l, \ Gq_l, \ Iq_l \quad (l = 1, 2, \ldots, m),$$

où le premier signe remplace l'adresse de la case considérée (cf. plus haut) et le deuxième met le bloc logique dans l'état sous-jacent. Les signes D, G, I, $q_1, \ldots, q_m$ forment l'*alphabet intérieur* de la machine.

Le fait que le bloc logique sert également à former à chaque manipulation donnée la commande qui sera introduite au commencement de la manipulation suivante dans le bloc de commande est une particularité de la machine de Turing. Ainsi, le bloc logique possède, outre le canal de sortie du signe s_j, deux autres canaux pour la sortie des deux signes de la commande suivante (voir le schéma correspondant sur la figure 5). Il importe de souligner que le *triplet de sortie* formé des signes s_j, P, q_l*) ne dépend que du *couple d'entrée* des signes s_i, q_n introduit lors de la même manipulation dans le bloc. Cela signifie que le bloc logique réalise la fonction faisant correspondre à chaque couple de signes s_i, q_n (il y en a en tout $k \cdot m$) le triplet de signes s_j, P, q_l. Cette fonction, que nous appellerons *fonction logique* de la machine, peut être commodément représentée sous la

*) Par P on entend ici n'importe lequel de trois signes D, G, I.

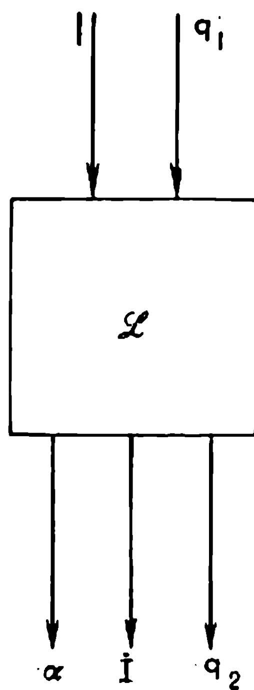


Fig. 5

	q_1	q_2	q_3	q_4	q_5
Λ	$\Lambda D q_2$	$\Lambda G q_3$	$\Lambda D q_1$	$\Lambda \dot{I} q_5$	$\Lambda \dot{I} q_5$
$\mid$	$\alpha \dot{I} q_2$	$\beta \dot{I} q_1$	$\mid D q_1$	$\mid G q_1$	$\mid \dot{I} q_5$
α	$\alpha G q_1$	$\alpha D q_2$	$\mid G q_3$	$\Lambda D q_4$	$\alpha \dot{I} q_5$
β	$\beta G q_1$	$\beta D q_2$	$\Lambda G q_3$	$\mid D q_4$	$\beta \dot{I} q_5$

Fig. 6

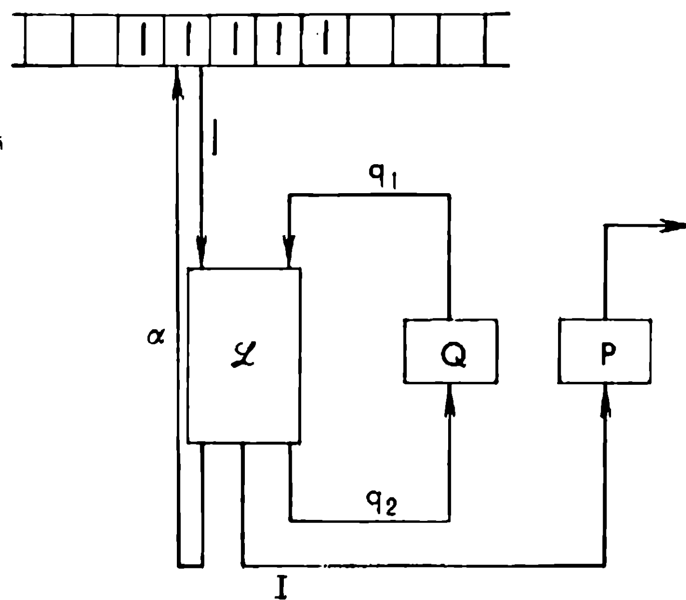


Fig. 7

forme d'un tableau rectangulaire dont les colonnes portent les numéros des signes des états et dont les lignes sont numérotées par les signes de l'alphabet extérieur ; dans chaque case de ce tableau est inscrit le triplet de sortie correspondant. Nous l'appellerons *schéma fonctionnel* de la machine ; un tel schéma est représenté sur la figure 6.

De ce que nous avons dit plus haut, il est clair que le *fonctionnement* de la machine de Turing est entièrement déterminé par la fonction logique réalisée dans le bloc logique. En d'autres termes, deux machines de Turing ayant le même schéma fonctionnel sont indiscernables, car nous ne nous intéressons qu'à leur fonctionnement. D'autre part, la structure de la machine, la constitution de ses organes isolés et leur interaction peuvent également être représentées sous forme d'un *schéma structural* commun pour toutes les machines de Turing (cf. fig. 7).

Dans ce schéma, on a représenté la division de la mémoire en mémoires extérieure et intérieure. La mémoire extérieure est représentée par les cases d'une bande infinie, destinées à stocker l'information codée par les symboles de l'alphabet extérieur ; la mémoire intérieure, par deux cases appelées à stocker la commande suivante : la Q case conserve le signe de l'état et la P case celui du décalage. Dans ces deux cases a lieu le retard des signes P, q_l obtenus à la sortie du bloc logique lors de la manipulation considérée jusqu'à moment où ils sont introduits dans le bloc de commande (manipulation suivante). Les fonctions du bloc de commande s'avèrent maintenant extrêmement simplifiées et consistent au fond à faire progresser la bande d'une case seulement, conformément au signe introduit P. On aurait pu en fait faire passer le signe d'état directement de la Q case dans $\mathcal{L}$, en formant ainsi une ligne, dite de réaction, suivant laquelle le signe q_l, formé dans le bloc $\mathcal{L}$ lors de la manipulation précédente, est introduit dans ce même bloc.

Le fonctionnement de la machine de Turing peut être décrit de la manière suivante. Avant sa mise en marche on

enregistre sur la bande une information initiale (suite de 5 traits sur la figure 7), et, dans le « champ visuel » de la machine, on fixe une case initiale quelconque (la case contenant le quatrième trait à gauche sur la figure 7). Dans les cases P et Q sont introduits les signes d'état et de décalage initiaux (disons q_1 et I). Le fonctionnement ultérieur s'effectue de façon automatique et est déterminé de manière bijective par le schéma fonctionnel de la machine. Considérons, par exemple, le cas où l'on se donne le schéma fonctionnel de la figure 6.

● PREMIÈRE PÉRIODE. Le signe visualisé est | (trait) dans la case initiale (décalage I), à l'état q_1. Résultat : le triplet de sortie est αIq_2, c'est-à-dire le signe | est remplacé par α, et dans les cases P et Q est stockée la commande Iq_2 jusqu'à la période suivante.

● DEUXIÈME PÉRIODE. Le signe visualisé est α de la même case (décalage I) à l'état q_2. Triplet de sortie : αDq_2, c'est-à-dire le signe α reste inchangé lors du passage à la commande Dq_2.

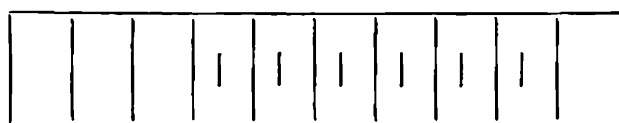
● TROISIÈME PÉRIODE. Le signe visualisé est le trait qui figure dans la case voisine à droite (décalage D) à l'état q_2. Résultat : le signe | est remplacé par β avec le passage à la commande Iq_1, etc.

Comme l'on voit sur la dernière colonne du schéma fonctionnel (fig. 6) la machine ne s'arrête qu'avec l'apparition de l'état q_5 à un certain stade du processus. En effet, quel que soit le signe examiné, il ne sera remplacé par aucun autre et la machine continuera de l'examiner (décalage I) dans le même état q_5. Cet état est l'*état d'arrêt* qui marque la fin du processus de calcul dans le cas où la machine est *applicable* à l'information introduite avant sa mise en marche.

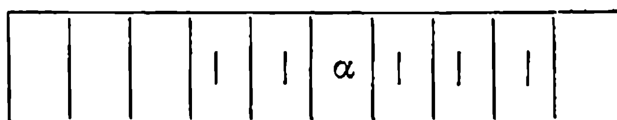
Un tel schéma fonctionnel peut au fond être utilisé par l'homme ; il représente alors un certain procédé standard de

donner un algorithme de transformation des données initiales, inscrites dans l'alphabet extérieur, en résultat correspondant, inscrit dans le même alphabet. C'est ce que nous venons de faire avec le schéma fonctionnel de la figure 6 en transformant le mot formé de 5 traits, tout en reconnaissant que la machine de Turing aurait fait la même chose. Pour rendre notre exposé plus clair, nous utiliserons dans la suite la notion de *configurations.* Par k-ième *configuration* on entend l'image de la bande de la machine contenant l'information qu'on enregistre sur elle au début de la k-ième manipulation ; de plus, dans la case examinée, on écrit le signe de l'état qui est introduit dans le bloc logique $\mathscr{L}$ au début de cette manipulation. Ainsi, dans la k-ième configuration, est indiqué de façon explicite le couple de signes d'entrée ; on peut donc, en s'adressant au schéma fonctionnel, déterminer le triplet de sortie correspondant et, par là même, la $(k + 1)$-ième configuration.

Les première et deuxième configurations de l'exemple ci-dessus sont de la forme :



q_1



q_2

avec les couples d'entrée $1q_1$, αq_2 respectivement. Le passage de la première configuration à la deuxième est conditionné par le triplet de sortie $\alpha I q_2$ qui correspond sur le schéma de la figure 6 au couple d'entrée $1q_1$.

Convenons encore de simplifier quelque peu l'écriture des schémas fonctionnels pour les rendre plus accessibles

	q_1	q_2	q_3	q_4
Λ	Dq_3	Gq_3	Dq_1	!
I	αq_2	Gq_3	Dq_1	!
α	G	D	I G	Λ D
β	G	D	Λ G	I D

Fig. 8

à examiner et plus commodes pour écrire les configurations. Nous refuserons d'écrire complètement le triplet de sortie s_jPq_l, en omettant les signes s_j et q_l, s'ils sont identiques aux signes d'entrée correspondants, ainsi que le signe I qui marque l'absence de décalage. Cela nous permet, en particulier, d'omettre complètement la colonne correspondant à l'état d'arrêt. La notation simplifiée du schéma de la figure 6 est donnée sur la figure 8, dans laquelle l'état d'arrêt est désigné par le signe «!». A partir de la colonne q_1 du schéma de la figure 8 il est plus aisé de voir sur la figure 6 que la machine, qui se trouve dans l'état q_1 avec le signe examiné α, commence une série de décalages vers la gauche à travers tous les α et β qui se suivent tout en restant dans l'état q_1 et sans modifier le contenu des cases examinées jusqu'à ce que dans son champ visuel apparaisse pour la première fois un trait ou une case vide ; c'est seulement dans ces conditions que la machine sort de l'état q_1.

Le signe «! » sera utilisé dans la suite pour marquer l'état d'arrêt de la machine.

§ 8. RÉALISATION D'UN ALGORITHME DANS LA MACHINE DE TURING

En faisant appel à une série d'exemples, nous montrerons dans ce paragraphe comment on construit les machines de Turing qui réalisent certains algorithmes arithmétiques

simples, et comment se développe dans la machine le processus de réalisation de ces algorithmes. Par construction d'une machine on entend, conformément au contenu du paragraphe précédent, la formation du schéma fonctionnel de son bloc logique, qui représente une certaine forme standard d'écriture de l'algorithme. Dans ce même paragraphe nous exposerons également certaines considérations d'ordre général sur la méthode de construction des machines de Turing (des schémas fonctionnels).

I. Algoritme de passage de n à $n+1$ dans le système décimal

On résout le problème du type suivant:

Etant donnée l'écriture décimale d'un nombre n (c'est-à-dire la représentation d'un entier naturel n dans le système décimal) *on cherche à donner l'écriture décimale du nombre* $n+1$.

On prend pour cela l'alphabet extérieur composé de dix chiffres 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 et du signe vide $\wedge$. La machine ne peut être que dans deux états: q_0 (état de fonctionnement) et ! (arrêt). Le nombre donné n, et celui qui en résulte $n+1$ seront écrits dans le système décimal et leurs chiffres occuperont chacun une case (les cases se suivent dans l'ordre, sans case vide). Le schéma fonctionnel correspondant est représenté sur la figure 9 sous forme de la partie du tableau que l'on obtient en ne tenant pas compte de la dernière ligne ni de la dernière colonne (le sens du tableau complété sera donné plus loin). Admettons qu'au moment de sa mise en marche la machine lit le chiffre du rang des unités du nombre n et qu'elle se trouve dans l'état q_0 ; si ce chiffre est différant de 9, la machine s'arrêtera déjà après la première période, au cours de laquelle le chiffre considéré est remplacé par un autre en conformité avec le schéma. Si le dernier chiffre est 9, la machine le remplace par 0 et effectue un décalage à gauche (vers le rang voisin

d'ordre plus grand) ; elle reste dans l'état de fonctionnement q_0 (on assure ainsi le transfert d'une unité dans les rangs plus élevés). Si le nombre se termine par $k9$, la machine s'arrêtera après avoir effectué exactement la $k + 1$ période. Sur la figure 10 sont représentées les configurations correspondantes pour $n = 389$.

Elucidons maintenant le sens du tableau complété de la figure 9. Il définit le schéma fonctionnel de la machine où l'on a ajouté encore un état q_1 ; de plus, dans l'alphabet extérieur figure le signe supplémentaire, le « trait ». Si au moment de sa mise en marche la machine est réglée sur l'état q_0 et s'il n'y a pas un seul « trait » sur la bande, le processus sera absolument identique à celui décrit plus haut. Ceci découle immédiatement du fait que, dans ces conditions, la dernière ligne et la dernière colonne n'interviennent pas dans le processus. Cela signifie, en particulier, que cette dernière machine peut aussi être utilisée pour la réalisation de l'algorithme précédent. Néanmoins ce n'est pas par hasard que nous l'examinons, car elle est capable d'effectuer certaines autres opérations.

Supposons que sur la bande soit enregistrée l'écriture décimale du nombre n et que dans certaines cases situées à droite de cette inscription soient écrits des traits, à raison de un par case. Examinons le fonctionnement de la machine suivant ce schéma, en considérant qu'avant le départ elle lit le trait le plus à droite et qu'elle est réglée sur l'état q_1. Lors de la première manipulation (couple d'entrée q_1 |), ce trait est effacé et la machine effectue un décalage à gauche avec passage à l'état q_0 (triplet de sortie $\wedge\, Gq_0$). Lors des périodes ultérieures se poursuit (toujours dans l'état q_0) un décalage à gauche à travers tous les traits jusqu'au premier chiffre du rang des unités. A partir de ce moment-là le processus s'avère parfaitement identique à celui de l'algorithme précédent, c'est-à-dire l'écriture du nombre n se transforme en celle du nombre $n + 1$, et le processus est réalisé.

	q_0	q_1
0	1 !	
1	2 !	
2	3 !	
3	4 !	
4	5 !	
5	6 !	
6	7 !	
7	8 !	
8	9 !	
9	0 G	
Λ	1 !	
I	G	Λ G q_0

Fig. 9

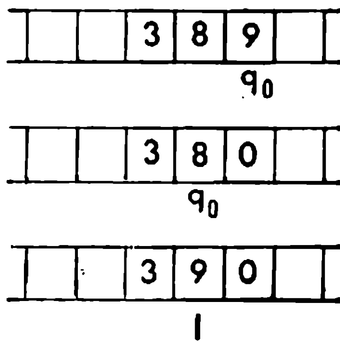


Fig. 10

En bref, la machine fait diminuer d'une unité le nombre de traits et réalise, dans le système décimal, le passage du nombre n au nombre $n + 1$. Convenons d'appeler ce processus *passage contrôlé* de l'écriture décimale du nombre n à celle du nombre $n + 1$.

Sur la figure 11 sont représentées les configurations pour l'ensemble de 5 traits et pour $n = 389$.

II. Algorithme de passage au système décimal

Construisons le schéma fonctionnel de la machine (algorithme) de résolution d'un problème du type suivant:

Etant donné un ensemble fini des traits inscrits dans des cases, à la suite les uns des autres et sans omission (convenons d'appeler de tels ensembles *collections de traits*) ; *on demande de noter dans le système décimal le nombre de ces traits.*

Plus brièvement: compter ces traits.

Un tel schéma est représenté sur la figure 12. Pour se convaincre que ce schéma décrit effectivement l'algorithme cherché il est utile de le comparer avec celui représenté sur la figure 9. La seule différence en est que dans la colonne q_0 du schéma de la figure 12, au lieu de l'état «! » (schéma de la figure 9), figure un nouvel état q_2 ; la différence des colonnes q_1 est sans importance pour le schéma de la figure 9. C'est pourquoi, si sur la bande sont enregistrées l'écriture décimale du nombre n et, à sa droite, la collection de ses traits, et si, dans le champ visuel de la machine, on place comme précédemment le trait le plus à droite (la machine est toujours dans l'état q_1), alors le stade initial du processus sera parfaitement identique à celui du schéma de la figure 9: le trait de la collection sera effacé et l'écriture du nombre n sera remplacée par celle du nombre $n + 1$. Mais au moment où, selon le schéma de la figure 9, à ce stade du processus apparaît l'état! (arrêt du processus), selon le schéma de la figure 12, apparaît l'état q_2, et le processus se poursuit. Si, en particulier, la première configuration est

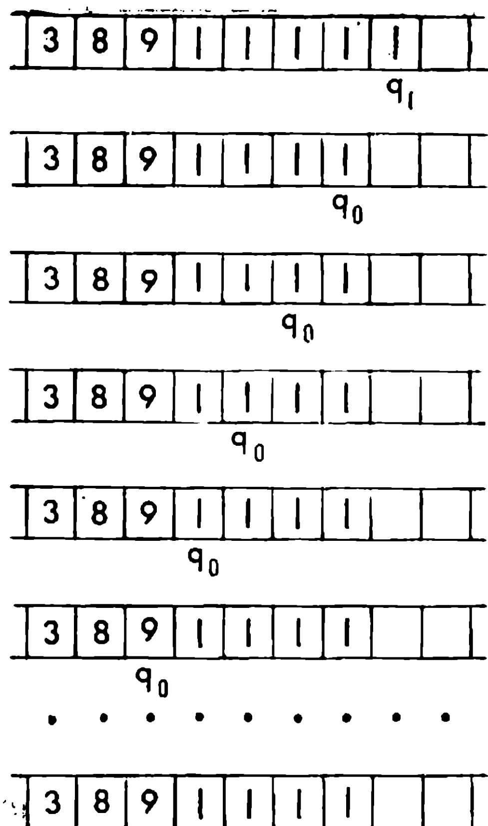


Fig. 11

	q_0	q_1	q_2
0	1 q_2	!	D
1	2 q_2	!	D
2	3 q_2	!	D
3	4 q_2	!	D
4	5 q_2	!	D
5	6 q_2	!	D
6	7 q_2	!	D
7	8 q_2	!	D
8	9 q_2	!	D
9	0 G	!	D
Λ	1 q_2	!	G q_1
|	G	Λ G q_0	D

Fig. 12

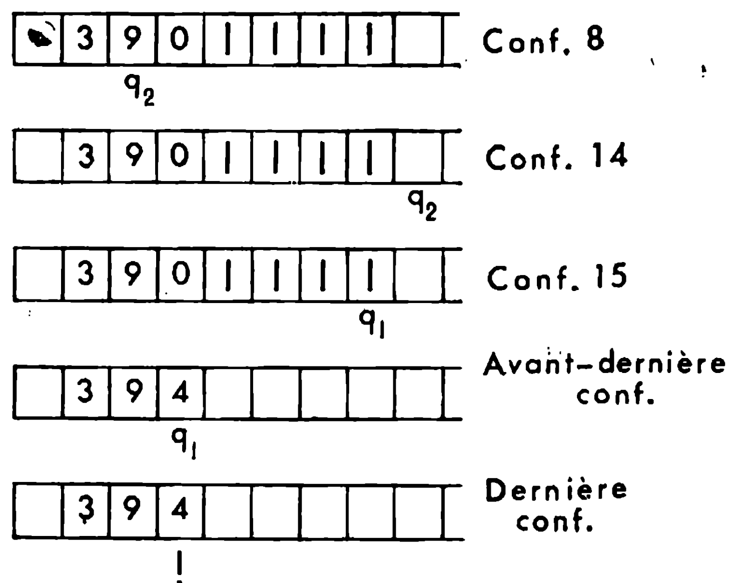


Fig. 13

celle représentée sur la figure 11, la huitième prend alors la forme de la figure 13. La colonne de l'état q_2 permet de suivre la marche du processus ultérieur : une série de décalages à droite s'amorce à travers tous les chiffres et tous les traits, jusqu'à l'apparition de la première case vide pour le couple d'entrée $\wedge q_2$ (cf. la configuration 14 de la figure 13) ; puis vient un décalage à gauche avec passage simultané à l'état q_1 (configuration 15 de la figure 13). Ainsi, dans le champ visuel de la machine apparaît de nouveau le trait le plus à droite de la collection dans l'état q_1 ; c'est par là que se termine un cycle du fonctionnement et que commence un second analogue au premier. Lors du deuxième cycle, un trait encore est effacé, et l'écriture du nombre $n + 1$ est remplacée par celle du nombre $n + 2$. Si l'on avait initialement k traits dans la collection ils seront tous effacés après k cycles et l'écriture initiale du nombre n sera remplacée par celle du nombre $n + k$. Après le k-ième cycle la machine se trouvera à nouveau dans l'état q_1 mais dans son champ visuel se trouvera cette fois non pas un trait (ceux-ci étant tous effacés) mais le premier chiffre (i.e. le chiffre des unités) dans l'écriture du nombre $n + k$ (avant-dernière configuration sur la figure 13). Comme on le voit sur le schéma de la figure 12, le processus s'arrêtera dans ce cas (dernière configuration sur la figure 13).

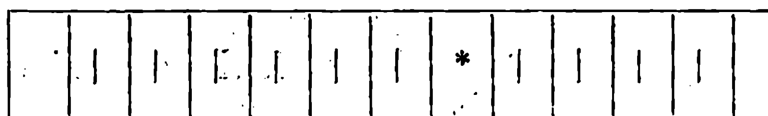
De tout ce qu'on a dit plus haut résulte que, si au moment de la mise en marche de la machine sont enregistrés sur la bande le chiffre 0 et une collection de k traits, la machine effacera tous ces traits et remplacera 0 par l'écriture décimale du nombre $0 + k$, c'est-à-dire k. En fait, on peut se passer de 0 au début des opérations, car si le signe $\wedge$ figure à la place de 0, dans les états q_0 et q_1 la machine se comporte de la même façon que si l'on avait 0 (cf. schéma de la figure 12). Ainsi, le schéma proposé à la figure 12 décrit réellement l'algorithme de passage de la collection des traits à l'écriture décimale du nombre d'éléments dans cette collection.

● EXERCICE. Etablir par analogie avec I le schéma fonctionnel de la machine (algorithme) qui réalise le passage de l'écriture décimale par nombre n à celle du nombre $n - 1$ (pour $n \geqslant 1$). Ensuite, par analogie avec II, établir le schéma fonctionnel de la machine pour la traduction du système décimal de calcul, i.e. pour le remplacement de l'écriture décimale d'un nombre n quelconque par une collection de n traits.

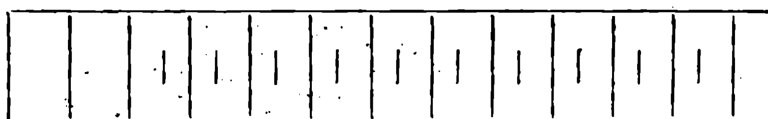
Considérons maintenant quelques exemples de machines de Turing pour la résolution de problèmes arithmétiques. Dans ces problèmes les données initiales (conditions du problème) aussi bien que le résultat obtenu sont des entiers naturels. Convenons de considérer que chaque entier naturel est donné par une collection formée du même nombre de traits; si dans le problème figurent plusieurs entiers naturels, alors les collections de traits qui les représentent seront marquées d'un signe particulier, par exemple d'un astérisque * qui fait aussi partie de l'alphabet extérieur de la machine.

III. Algorithme d'addition

Sur la bande est enregistré un couple de nombres, par exemple,



Il faut obtenir leur somme, c'est-à-dire



Il est à remarquer qu'on ne peut pas réduire cette opération à une simple suppression de l'astérisque, car alors apparaît une case vide entre les traits et, par conséquent, l'ensemble des traits restants ne peut plus être considéré comme représentant un entier naturel. Selon le schéma fonc-

	q_0	q_1	q_2
|	$\wedge$ D q_2	G	D
$\wedge$	D	D q_0	| q_1
•	$\wedge$!	• G	D

Fig. 14

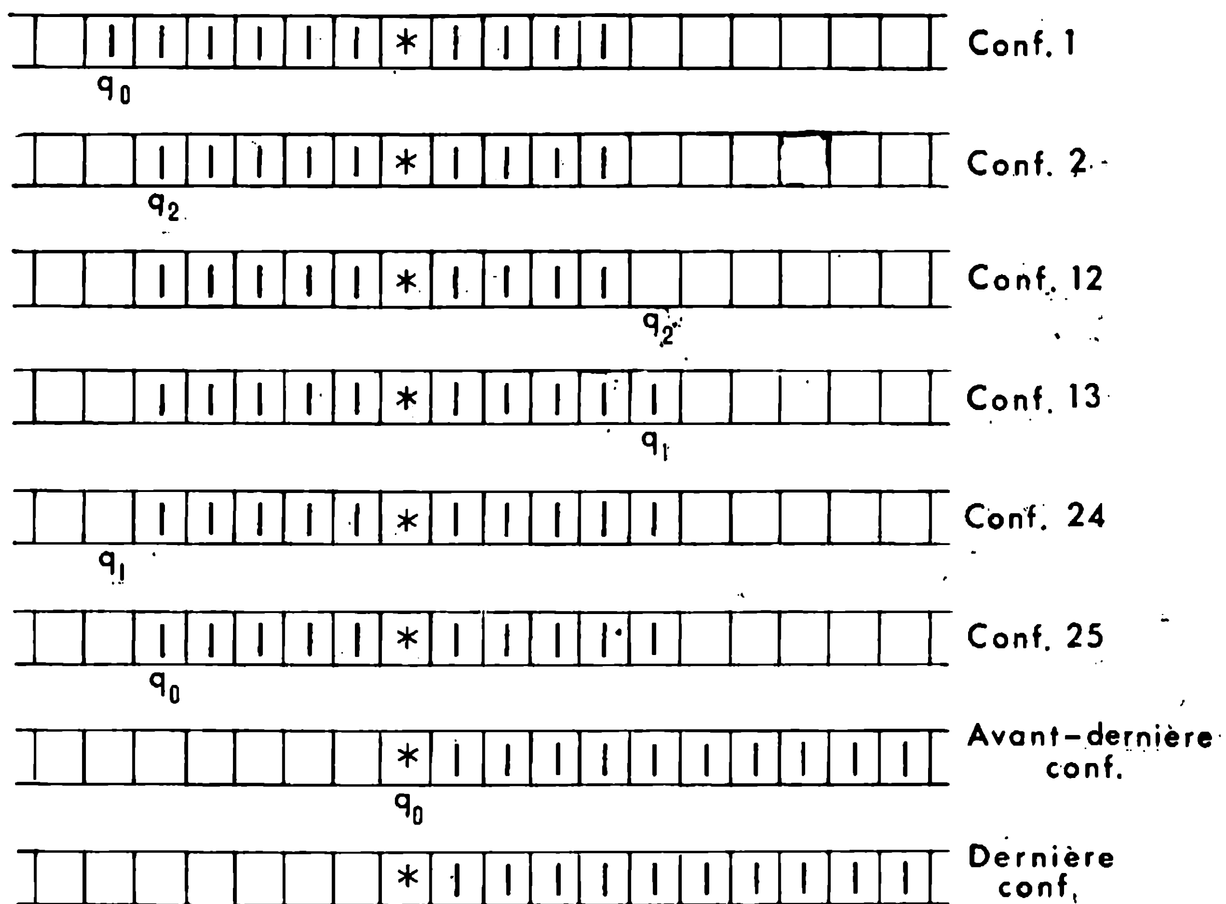


Fig. 15

tionnel proposé ci-dessous (cf. fig. 14) le fonctionnement de la machine est le suivant.

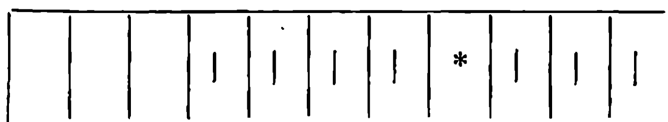
Conditions initiales : dans le champ visuel on fixe le trait le plus à gauche et la machine est dans l'état q_0 (configuration 1 de la figure 15).

● PREMIÈRE PÉRIODE. Le trait examiné est effacé, la machine effectue un décalage à droite (pour fixer dans le champ de vision le trait suivant) et passe à l'état q_2 (configuration 2).

Les périodes suivantes (comme on le voit sur la colonne q_2) se ramènent à une série de décalages à droite par tous les traits et l'astérisque, jusqu'à l'apparition de la première case vide (configuration 12) ; alors un trait s'inscrit dans cette case vide (couple d'entrée $\wedge q_2$) et la machine passe à l'état q_1 (configuration 13). A l'état q_1, une série de décalages ont lieu vers la gauche par tous les traits et l'astérisque, jusqu'à l'apparition de la première case vide à gauche (configuration 24) ; alors on a un décalage à droite (couple d'entrée $\wedge q_1$) et le premier des traits restants, à gauche de l'astérisque, se met dans le champ visuel et la machine passe à l état q_0 (configuration 25). Après ce cycle, un trait du terme de gauche est passé dans le terme de droite. S'il y avait initialement à gauche de l'astérisque k traits, alors, après k cycles, ils seront tous passés à droite de l'astérisque. Lors du $(k + 1)$-ième décalage à droite, dans le champ visuel de la machine, à l'état q_0, se trouve non pas un trait (il n'y en a plus à gauche de l'astérisque), mais l'astérisque lui-même (avant-dernière configuration). L'astérisque (couple d'entrée * q_0) s'efface alors, et la machine s'arrête (dernière configuration). Ainsi, on a obtenu la somme cherchée.

IV. Algorithmes de sommation et miliplication itérées

Examinons comment il faut modifier le schéma de la figure 14 pour qu'après l'enregistrement initial sur la bande d'un couple de nombres m, n, par exemple :



la machine puisse en arriver à un processus illimité consistant à ajouter le nombre de gauche m à celui de droite, puis à ajouter m à la somme obtenue $n + m$, puis à la somme $n + 2m$, etc. Pour cela il faut, évidemment, que le terme de gauche ne disparaisse pas complètement après la première addition mais, au contraire, qu'il puisse être rétabli après chaque addition et être ajouté au nombre qui figure à droite de l'astérisque. Pour cela il suffit, par exemple, de ne pas effacer les traits de la collection gauche mais de les remplacer momentanément par certains signes ou annotations. Sur la figure 16, est représenté un schéma où la lettre α joue le rôle d'une telle annotation. Conformément à cela, l'inclusion du signe vide Λ dans la première ligne du schéma 14 correspond à celle du symbole α dans la première ligne du schéma 16 et, de plus, la quatrième ligne du schéma de la figure 16 admet également l'inclusion de α ; les trois premières cases de cette ligne sont parfaitement identiques à celles de la ligne renfermant Λ sur la figure 14.

Pour que le processus ne s'arrête pas après la première addition, il faut remplacer le signe «! » qui apparaît sur la figure 14 pour le couple d'entrée * q_0 par le signe d'un autre état assurant la continuation du processus ; dans le cas considéré on a introduit l'état supplémentaire q_3 et on lui a réservé une colonne à part.

Dans le schéma de la figure 16 ne figure pas le signe «! » (arrêt) ; le processus que ce schéma décrit est donc illimité. Nous laissons au lecteur le soin de vérifier qu'une telle machine réalise l'addition du nombre figurant à gauche de l'astérisque avec celui qui figure à sa droite et que ce processus est illimité. Si au début du travail il n'y a pas de traits à droite de l'astérisque (c'est-à-dire si le deuxième nombre est zéro), il y apparaît d'abord m traits à droite de l'astérisque, puis $2m$, $3m$ et ainsi de suite.

	q_0	q_1	q_2	q_3
I	αDq_2	G	D	
Λ	D	Dq_0	I q_1	Dq_0
*	q_3	G	D	G
α	D	Dq_0	I q_1	I G

Fig. 16

● EXERCICE. Composer le schéma fonctionnel de l'algorithme de multiplication.

Indication. Modifier le schéma précédent de sorte que le processus de sommation itérée ne soit pas illimité mais s'effectue autant de fois qu'il y a d'unités dans le multiplicateur (après chaque cycle de sommation s'efface un trait dans la collection représentant le multiplicateur).

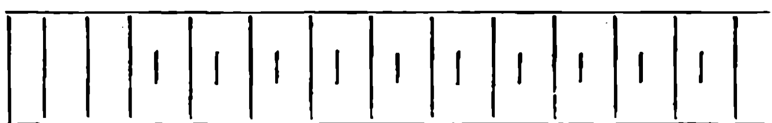
V. Algorithme d'Euclide

Considérons maintenant la réalisation sur machine de Turing de l'algorithme d'Euclide consistant à déterminer le P.G.C.D. des nombres a, b. Nous l'avons déjà décrit deux fois : une première fois, sous la forme d'une instruction verbale et la seconde, sous la forme d'un programme pour une machine électronique à commande automatique. Maintenant nous donnerons cet algorithme sous la forme d'un schéma fonctionnel pour la machine de Turing et examinerons le processus de calcul dans la machine. Ce processus est formé de *cycles de comparaison* et de *cycles de soustraction* alternatifs correspondant aux *opérations élémentaires de comparaison* et *de soustraction*, réalisées dans une machine électronique. Le schéma fonctionnel correspondant est représenté sur la figure 8 ; son alphabet extérieur est constitué de quatre signes

$$\Lambda, \; |, \; \alpha, \; \beta.$$

Les entiers naturels seront toujours représentés par des collections correspondantes de traits. Pour simplifier quelque peu notre tâche et ne pas entrer dans les détails secondaires susceptibles de compliquer notre étude, convenons-nous de disposer sur la bande les collections de traits représentant les deux nombres donnés à la suite les uns des autres, sans vide et sans les séparer par des astérisques. Au début du processus la machine a dans son champ visuel le trait le plus à droite de la collection de traits du premier nombre. Après l'analyse détaillée qui sera faite plus bas, le lecteur pourra sans difficultés particulières et à titre d'exercice modifier le schéma fonctionnel de façon qu'il continue d'assurer un fonctionnement correct de la machine avec d'autres données initiales du problème (si, par exemple, les collections de traits sont séparées par un astérisque et si l'on a dans le champ visuel de la machine une case vide quelconque). Il est à souligner que les lettres α, β serviront de symboles momentanés que l'opérateur utilise (habituellement sous la forme d'apostrophes ou de coche) pour se souvenir de certains facteurs qui sont intervenus dans le processus de calcul.

Pour illustrer ce problème, nous faisons appel aux configurations obtenues pour $a = 4$, $b = 6$. La première configuration est de la forme:



q_1

Dans le cycle de comparaison interviennent seulement deux états: q_1, q_2; dans celui de soustraction interviennent q_3 et q_4.

Examinons ce processus en détail. La machine commence par comparer les nombres fixés sur la bande pour déterminer le plus grand. Elle opère tout comme un homme qui comparerait deux longues suites d'unités dont la totalité est

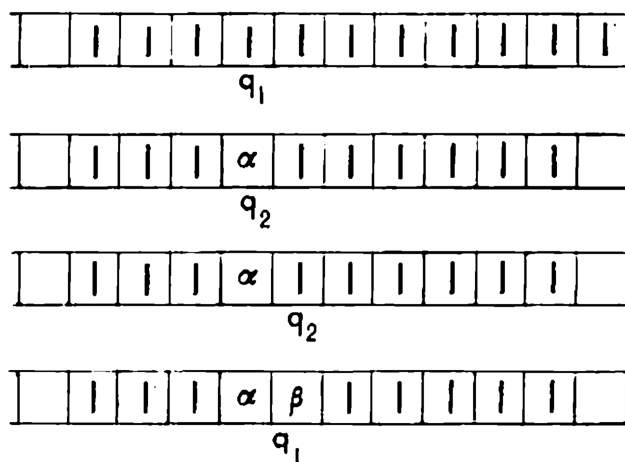


Fig. 17

difficile à embrasser d'un seul coup. Il marque à tour de rôle chaque unité dans les deux suites d'une manière ou d'une autre (par exemple, en les munissant d'apostrophes), lorsque l'une des deux suites disparaît, on voit quelle est la suite qui comporte le plus grand nombre d'unités.

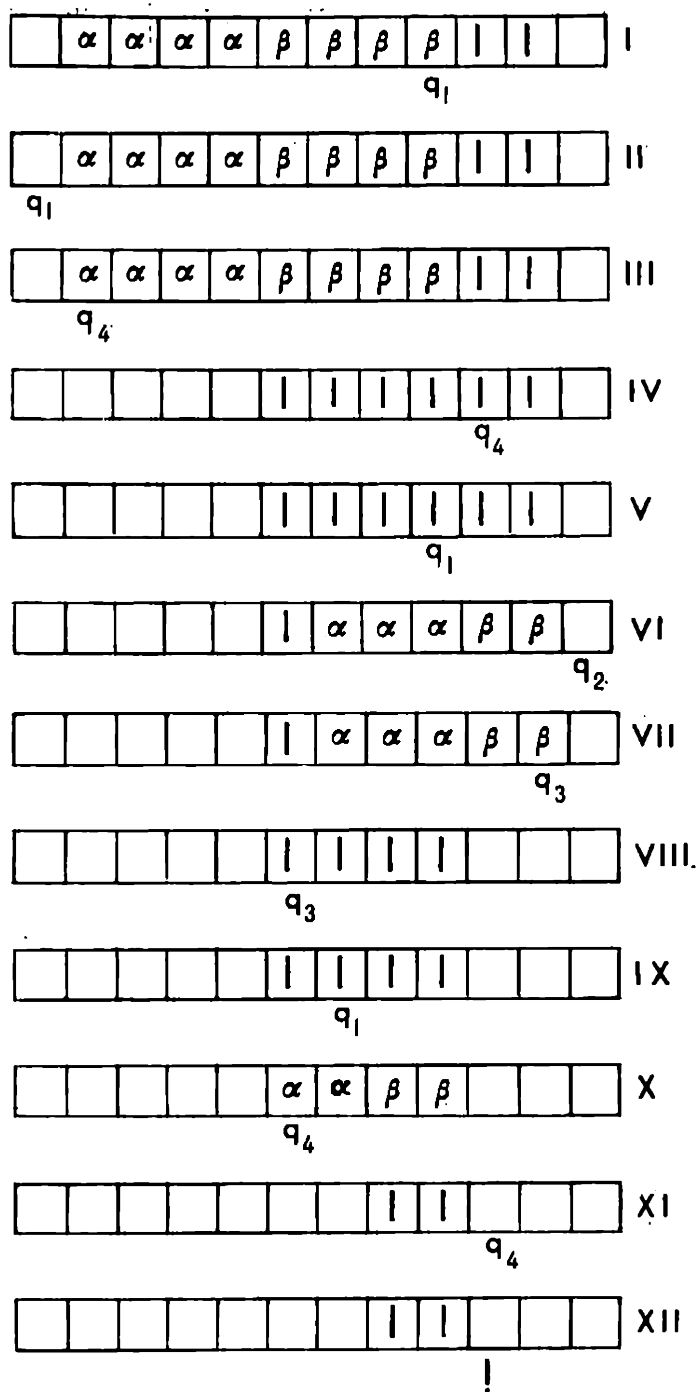
La machine remplace un trait du premier nombre par le symbole α, puis un trait du deuxième nombre par β ; elle revient ensuite aux traits du premier nombre pour remplacer l'un d'eux par α, puis remplace l'un des traits du deuxième nombre par β, etc.

Les configurations correspondant aux quatre premières périodes sont représentées sur la figure 17 ; vers la fin de la quatrième période la machine a eu le temps de marquer chaque trait des deux nombres et amorce alors un décalage à gauche pour trouver le trait le plus proche, non encore marqué du nombre de gauche. Après quelques manipulations encore, apparaît sur la bande la configuration I de la figure 18, i.e. le premier nombre est déjà épuisé alors que le deuxième ne l'est pas encore. La recherche d'un trait à gauche est vaine et conduit à la configuration II ; le cycle de comparaison est déjà réalisé dans les seuls états q_1 et q_2. La manipulation suivante amène déjà à la configuration III.

Comme on le voit sur la colonne représentant l'état q_4 sur le schéma de la figure 8, la machine amorce maintenant le décalage à droite en remplaçant tous les α par les signes vides Λ (c'est-à-dire en supprimant tous les α) et tous les β par les traits. Après que le dernier β de droite ait été remplacé par un trait, apparaît sur la bande la configuration IV de la figure 18, puis la configuration V. Ainsi, après le cycle de comparaison a eu lieu un cycle de soustraction du premier nombre du deuxième, à la suite duquel le plus petit nombre a est effacé, et le plus grand nombre b est partagé en a et $b - a$. La machine examine alors le dernier trait du premier de ces nombres et revient à l'état q_1. Cela signifie que le problème initial pour les nombres a, b se ramène au même problème pour les nombres a, $b - a$. C'est précisément sur ce principe qu'est basé, on le sait, l'algorithme d'Euclide.

Il va de soi que le cycle suivant est un cycle de comparaison. Mais cette fois-ci il s'achève sur la disparition du second des deux nombres (celui de droite) qui s'avère en l'occurrence le plus petit. Cela est visible après que la machine, qui a remplacé les trois traits du premier nombre, ne trouve plus de traits dans le deuxième, i.e. apparaît la configuration VI.

La période suivante donne lieu à la configuration VII. Au cours de celle-ci la machine commence par soustraire le deuxième nombre du premier, c'est-à-dire efface tous les β en remplaçant tous les α par des traits. Après le remplacement du dernier α (à gauche) par un trait apparaît sur la bande la configuration VIII, puis la configuration IX; là-dessus s'achève le cycle de soustraction et commence celui de comparaison, etc. Ce processus se poursuit jusqu'à ce que le problème soit ramené au cas de deux nombres égaux (c'est ce que l'on a déjà ici). Alors commence le dernier cycle de comparaison qui doit conduire au résultat final. En effet, une fois la configuration X obtenue, le processus de soustraction engendre la configuration XI puis la configuration finale XII.



α α α α β β β β | | I
q1
α α α α β β β β | | II
q1
α α α α β β β β | | III
q4
| | | | | | IV
q4
| | | | | | V
q1
| α α α β β VI
q2
| α α α β β VII
q3
| | | | VIII
q3
| | | | IX
q1
α α β β X
q4
| | XI
q4
| | XII
!

Fig. 18

VI. Combinaison des algorithmes

Lors de la construction de nouveaux schémas fonctionnels il s'avère souvent commode de recourir aux schémas déjà construits ; cela est possible dans les cas où l'on considère un algorithme qui est dans un certain sens combinaison des algorithmes déjà étudiés. En voici un exemple. Soit à construire le schéma fonctionnel d'un algorithme transformant un couple de nombres a, b donnés par des collections correspondantes de traits en leur P.G.C.D. (écrit dans le système décimal). Comme cet algorithme peut être obtenu à la suite d'une *composition*, c'est-à-dire de l'application successive des deux algorithmes considérés plus haut: d'abord V, puis II, son schéma fonctionnel, représenté sur la figure 19, peut également être construit à l'aide d'une combinaison appropriée des schémas de la figure 8 et de la figure 12. Désignons par p_0, p_1, p_2 les états propres au schéma de la figure 12 pour les différencier des états du schéma de la figure 8 et par p_2 le symbole d'arrêt «! » de la figure 8. Réunissons, enfin, les schémas ainsi modifiés en un seul, comme il est indiqué sur la figure 19. (Les cases non utilisées correspondent à des couples d'entrée qui ne participent pas au processus examiné ; on peut donc y inscrire n'importe quel triplet de sortie.) Il est aisé maintenant de voir que, conformément au schéma de la figure 19, on aura d'abord un processus de transformation des deux collections données de traits jusqu'à ce qu'apparaisse sur la bande une collection représentant le P.G.C.D. des nombres considérés ; mais l'état d'arrêt ! (cf., par exemple, la configuration XII de la fig. 18) sera remplacé par l'état p_2 et le processus se poursuivra en continuant la transformation de cette collection en l'écriture décimale du P.G.C.D.

Il est facile de comprendre que cette méthode peut être généralisée à la composition d'un nombre fini d'algorithmes.

	q_1	q_2	q_3	q_4	p_0	p_1	p_2
0					1 p_2	!	D
1					2 p_2	!	D
2					3 p_2	!	D
3					4 p_2	!	D
4					5 p_2	!	D
5					6 p_2	!	D
6					7 p_2	!	D
7					8 p_2	!	D
8					9 p_2	!	D
9					0 G	!	D
Λ	Dq_4	Gq_3	Dq_1	p_2	1 p_2	!	Gp_1
$\mid$	αq_2	βq_1	Dq_1	Gq_1	G	ΛGp_0	D
α	G	D	$\mid$ G	Λ D			
β	G	D	Λ G	$\mid$ D			

Fig. 19

Il s'ensuit en particulier que les entiers naturels peuvent être considérés dans les algorithmes numériques comme donnés par des collections de traits, car il est aisé de passer d'un schéma $\mathfrak{A}$ construit dans ces conditions à un schéma $\mathfrak{B}$ dans le système décimal. Il suffit pour cela de construire le schéma $\mathfrak{B}$ d'après la méthode décrite ci-dessus par composition des schémas de trois algorithmes : l'algorithme du pas-

sage du système décimal dans un autre, l'algorithme de $\mathfrak{A}$ et, enfin, celui du passage dans le système décimal *).

Un autre procédé de combinaison des algorithmes consiste à appliquer itérativement un seul et même algorithme jusqu'à ce que soit remplie la condition indiquée à l'avance. Ainsi, par exemple, l'algorithme du passage dans le système décimal se ramène à l'application itérée de l'algorithme du passage contrôlé de n à $n+1$ jusqu'à ce qu'il ne reste aucun trait. Le schéma fonctionnel de cet algorithme (si on l'a déjà) et la condition avancée permettent de construire le schéma de l'algorithme itéré, mais cette méthode est plus compliquée que celle de composition et nous ne l'examinerons pas ici en détail.

Les exemples examinés ci-dessus expliquent assez clairement comment composer les schémas fonctionnels pour d'autres algorithmes numériques, en particulier pour les algorithmes non numériques. Indiquons le plan général de composition du schéma fonctionnel pour l'algorithme de réduction des mots (cf. exemple 4 du § 3). On commence par former les schémas $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{A}_3$, $\mathfrak{A}_4$ qui réalisent les substitutions orientées

$$b \to acc$$

$$ca \to accc$$

$$aa \to \Lambda$$

$$cccc \to \Lambda$$

respectivement. Ainsi, par exemple, la machine qui a le schéma $\mathfrak{A}_4$ transforme tout mot de l'alphabet $\{a, b, c\}$

*) L'analogie avec les machines à calculer électroniques fonctionnant dans le système binaire vient naturellement à l'esprit: celles-ci possèdent un mécanisme transformant les données initiales du système décimal en système binaire et un autre, qui transforme de nouveau le résultat final en système décimal.

inscrit sur la bande en un mot qui s'obtient en suppriman le quadruplet de c consécutifs, situé le plus à gauche ; s'il n'y en a pas, le mot reste alors inchangé. On construit ensuite les schémas $\widetilde{\mathfrak{A}}_1$, $\widetilde{\mathfrak{A}}_2$, $\widetilde{\mathfrak{A}}_3$, $\widetilde{\mathfrak{A}}_4$ pour répéter les algorithmes $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{A}_3$, $\mathfrak{A}_4$ respectivement. Ainsi, par exemple, la machine qui a le schéma $\widetilde{\mathfrak{A}}_4$ supprime d'abord un premier quadruplet de c, puis un autre, etc., jusqu'à ce que l'on obtienne enfin un mot où il n'y a plus de quadruplet de c consécutifs. Enfin, le schéma cherché pour l'algorithme de réduction représente une composition des schémas $\widetilde{\mathfrak{A}}_1$, $\widetilde{\mathfrak{A}}_2$, $\widetilde{\mathfrak{A}}_3$, $\widetilde{\mathfrak{A}}_4$.

§ 9. HYPOTHÈSE FONDAMENTALE DE LA THÉORIE DES ALGORITHMES

Le processus se déroulant au sein d'une machine de Turing rappelle en quelque sorte, à en juger par les exemples ci-dessus, le tournage cinématographique extrêmement ralenti d'un processus de calcul effectué par un opérateur, conformément à un certain algorithme. Ces mêmes exemples suggèrent également l'idée de donner, au moyen de schémas fonctionnels de machines de Turing, d'autres algorithmes connus, qui sont d'habitude donnés autrement, sous forme d'une instruction écrite ou de formules spéciales quelconques par exemple. Déjà à ce stade de nos raisonnements il semble d'ailleurs fort probable que cela puisse réussir aussi dans d'autres cas. En est-il ainsi ? A quel point peut-on généraliser les notions de machine de Turing et de schéma fonctionnel ? Peut-on enfin considérer comme universelle la donnée des algorithmes par des schémas fonctionnels ou, plus précisément, tout algorithme peut-il être construit de cette façon ?

La réponse à ces questions proposée par la théorie moderne des algorithmes peut être énoncée comme suit :

● **Hypothèse fondamentale de la théorie des algorithmes.** *Tout algorithme peut être donné à l'aide d'un schéma fonctionnel de Turing et réalisé dans la machine de Turing correspondante.*

A ce propos, nous examinerons les deux questions suivantes :

1. Quelle est la portée de cette hypothèse pour la théorie des algorithmes ?

2. Comment est-elle justifiée ?

Signalons avant tout la particularité suivante de cette formulation. Il s'agit dans cette formulation, d'une part, de tout algorithme, c'est-à-dire de la notion générale d'algorithme qui n'est pas, nous l'avons souligné à maintes reprises, une notion mathématiquement rigoureuse et, d'autre part, de la notion parfaitement précise de schéma fonctionnel de Turing. L'importance de l'hypothèse ci-dessus consiste donc en ce qu'elle précise la notion générale, mais vague, de « tout algorithme » à l'aide de la notion mathématique rigoureuse de schéma fonctionnel de Turing (et de sa réalisation dans une machine de Turing). Ainsi la théorie des algorithmes prend pour objet de recherche tous les schémas fonctionnels imaginables (machines de Turing). La question de savoir s'il est possible ou non de créer un algorithme pour la résolution des problèmes d'un type quelconque devient ainsi parfaitement légitime et se ramène au problème de l'existence ou non d'une machine de Turing (schéma fonctionnel) ayant la propriété requise.

Ainsi, l'hypothèse formulée rend légitime la définition fondamentale de la théorie moderne des algorithmes selon laquelle la notion vague d'algorithme s'identifie avec la notion rigoureuse de schéma fonctionnel d'une machine de Turing.

Où est donc la justification de cette hypothèse importante ? Il faut remarquer d'abord qu'il ne peut pas être question de la démontrer comme on le fait pour les théorèmes en mathématiques. En effet, cette formulation ne porte pas le

caractère d'un théorème, puisque c'est l'affirmation de la notion générale et non mathématiquement rigoureuse d'algorithme ; les raisonnements mathématiques rigoureux ne sont donc pas applicables.

La justesse de cette hypothèse repose surtout sur l'expérience. Tous les algorithmes connus, qui furent inventés au cours des millénaires de l'histoire des mathématiques, peuvent être donnés au moyen de schémas fonctionnels de Turing. Certes, le contenu de cette hypothèse n'est pas seulement tourné vers le passé et ne se limite pas à la constation de ce que tous les algorithmes connus peuvent être donnés par des schémas fonctionnels. Il a le caractère d'une prédiction: chaque fois qu'une instruction sera dans l'avenir reconnue comme un algorithme, on pourra toujours la donner sous la forme d'un schéma fonctionnel de Turing, quels que soient la forme et les moyens par lesquels cette instruction a été initialement donnée.

En ce sens, l'hypothèse fondamentale peut être comparée à une loi physique telle la loi de conservation de l'énergie sur laquelle s'appuient certains pronostics concernant l'avenir. C'est justement l'expérience pratique immense du passé qui sert de justification à de tels pronostics.

Mais il y a également d'autres considérations à l'appui de cette hypothèse.

Au paragraphe précédent nous avons signalé deux procédés servant à composer des algorithmes plus complexes à partir des algorithmes donnés: composition et itération des algorithmes. De tels procédés sont assez nombreux. Toutefois, tous ces procédés connus, et tous ceux que l'on peut prévoir dans l'état actuel de la science, permettent, lorsqu'on peut donner les algorithmes initiaux sous forme de schémas fonctionnels, d'en faire autant pour les algorithmes plus complexes qui en résultent. En particulier, pour la composition des algorithmes, nous avons déjà montré comment, à partir de schémas fonctionnels donnés,

on peut construire un nouveau schéma fonctionnel. Nous rappellerons en outre le fait suivant que nous avons mentionné au passage au § 6. Au moment où les mathématiciens se sont vus obligés d'élaborer la notion rigoureuse d'algorithme, ils ont entrepris des recherches afin d'établir une certaine forme standard de la donnée des algorithmes, suffisamment précise pour devenir un objet d'études mathématiques et suffisamment générale pour pouvoir y inclure tous les algorithmes imaginables. Outre les schémas fonctionnels des machines de Turing on a proposé d'autres moyens pour préciser cette notion. Ainsi, par exemple, A. Markov est arrivé à la notion d'*algorithme normal* (que nous avons illustrée en passant sur l'exemple de l'algorithme de réduction pour le calcul associatif de l'ex. 4 du § 3) ; et Hédel et Klini ont formulé la notion d'*algorithme récursif* (*fonction récursive*), etc. Néanmoins, toutes ces mises au point se sont avérées par la suite identiques. Il est à remarquer que ce fait est loin d'être fortuit ; il représente un argument de plus en faveur de l'hypothèse formulée.

Remarquons, pour conclure, qu'à l'intérieur de la théorie même des algorithmes, on n'applique pas l'hypothèse fondamentale, i.e. on ne fait pas appel à l'hypothèse fondamentale pour démontrer les théorèmes de cette théorie. Ainsi, même sans la connaître ou sans considérer comme valables les arguments apportés en sa faveur, on n'éprouvera pas de difficultés formelles dans la théorie moderne des algorithmes. Mais dans ce cas, ce que nous appelons *théorie des algorithmes*, sera seulement une théorie des schémas fonctionnels des machines de Turing, i.e., au fond, une théorie de certains algorithmes d'un type particulier.

L'auteur de ces lignes n'a jamais douté de la justesse de l'hypothèse fondamentale tout en considérant la théorie moderne des algorithmes qui en résulte comme une théorie générale décrivant la nature même des choses et non comme une théorie artificielle appelée à décrire des « algorithmes de Turing » particuliers.

§ 10. MACHINE UNIVERSELLE DE TURING

Jusqu'ici nous avons considéré que des algorithmes différents se réalisent dans des machines de Turing également différentes, qui se distinguent par leurs schémas fonctionnels. Néanmoins on peut construire une machine de Turing *universelle*, capable de réaliser dans un certain sens tout algorithme et donc d'effectuer le travail de toute machine de Turing.

Pour mieux assimiler cette idée, imaginons-nous l'expérience suivante. Supposons que soit enregistrée sur le ruban de la machine une information initiale $\mathfrak{A}$ et que l'on demande d'indiquer comment la machine transformera cette information et en quoi elle la transformera finalement. Si l'on est déjà familiarisé avec les principes du fonctionnement de telles machines il suffit de connaître, outre l'information initiale $\mathfrak{A}$, le schéma fonctionnel de cette machine. En imitant le fonctionnement de la machine et en notant les configurations nécessaires, tout comme nous l'avons fait en examinant l'algorithme d'Euclide, un opérateur obtiendra alors le même résultat que la machine. Mais cela signifie justement qu'un opérateur peut faire le travail de toute machine de Turing, à condition qu'il connaisse son schéma fonctionnel. Le processus même d'imitation de la machine conformément à son schéma fonctionnel peut être réglementé sous forme d'une instruction rigoureuse que peut réaliser tout homme n'ayant même pas la plus petite idée de machines de Turing. Si un opérateur dispose d'une telle instruction, qu'il est naturel d'appeler *algorithme d'imitation*, et se voit donner un schéma fonctionnel d'une machine de Turing et d'une information initiale quelconque enregistrée sur le ruban, il est alors capable d'imiter exactement le fonctionnement de la machine considérée et de produire finalement le même résultat. Un tel algorithme

d'imitation pourrait être donné sous la forme du système d'instructions suivant:

Instruction 1. Examiner sur le ruban la case (unique) sous laquelle a été écrite la lettre.

Instruction 2. Rechercher dans le tableau *) la colonne désignée par la même lettre qui se trouve sous la case considérée.

Instruction 3. Examiner dans cette colonne le triplet de lettres se trouvant à l'intersection de celle-ci avec la ligne portant la même lettre que la case examinée.

Instruction 4. Remplacer la lettre de la case considérée par la première lettre de ce triplet.

Instruction 5. Si la deuxième lettre du triplet est !, arrêter. Le processus est terminé.

Instruction 6. Si la deuxième lettre du triplet est *I*, remplacer la lettre écrite sous la case examinée par la troisième lettre du triplet.

Instruction 7. Si la deuxième lettre du triplet est *G*, supprimer alors la lettre sous la case examinée pour mettre à sa gauche la troisième lettre du triplet.

Instruction 8. Si la deuxième lettre du triplet est *D*, supprimer la lettre sous la case examinée pour inscrire à sa droite la troisième lettre du triplet.

Instruction 9. Revenir à l'instruction 1.

Il s'avère qu'une personne opérant conformément à cet algorithme peut être remplacée par une machine de Turing. Ce sera précisément une machine universelle capable d'imiter le travail de toute autre machine de Turing. En d'autres termes, cela signifie que l'algorithme d'imitation décrit plus haut au moyen d'un système de neuf instructions peut être donné implicitement sous la forme d'un schéma fonctionnel de Turing (schéma universel). Une démonstration exhaustive et rigoureuse de ce fait, servant d'ailleurs d'argument supplémentaire à l'appui de l'hypothèse fondamentale de la

*) C'est-à-dire dans le schéma fonctionnel.

théorie des algorithmes, est trop encombrante pour s'inscrire dans le cadre de ce petit ouvrage. Nous nous bornerons à quelques indications générales qui seront, semble-t-il, suffisantes pour comprendre le fond du problème.

Remarquons pour commencer que l'algorithme d'imitation que nous avons décrit a pour données initiales (information initiale) le schéma fonctionnel de la machine imitée et la configuration initiale correspondante. Cette information initiale est transformée par l'algorithme en une configuration finale qui représente le résultat issu de la machine imitée. La machine universelle doit faire la même chose. Mais ici il faut tenir compte des deux circonstances suivantes :

1. La transmission immédiate du schéma fonctionnel à la machine imitée et de la configuration correspondante sur le ruban de la machine universelle s'avère impossible en tant qu'information initiale. En effet, dans la machine universelle, comme dans toute machine de Turing, l'information est représentée par des lettres disposées en une dimension sur le ruban, c'est-à-dire sur une seule ligne, qui forme un ou plusieurs mots de l'alphabet extérieur de la machine. Mais tous les schémas fonctionnels examinés jusqu'ici ont été donnés au moyen de tableaux « bidimensionnels », c'est-à-dire de tableaux dans lesquels les lettres occupent plusieurs lignes. Il en est de même pour les configurations où les lettres marquant les états sont placées sous les lettres de l'alphabet extérieur (sous le ruban).

2. La machine universelle (comme toute machine de Turing) ne peut posséder qu'un alphabet extérieur fini déterminé. Mais elle doit pouvoir prendre comme information initiale tous les schémas et configurations possibles, dans lesquels on puisse rencontrer des lettres de différents alphabets avec un nombre arbitrairement grand de lettres différentes.

Il importe donc d'élaborer un procédé pour se donner les schémas fonctionnels et les configurations qui répondrait

à ces particularités propres à toute machine de Turing prise isolément (et surtout au caractère unidimensionnel de l'information et au caractère fini de l'alphabet). Venons-en à la description d'un tel procédé :

1. Au lieu de représenter un schéma sous forme d'un tableau bidimensionnel à k lignes et m colonnes, écrivons en suivant les mk quintuplets de lettres de ce tableau de sorte que le premier symbole du quintuplet indique la colonne du tableau, le second, la ligne, et les trois derniers représentent le triplet qui se trouve dans le tableau à l'intersection de la ligne et de la colonne considérées.

Ainsi, par exemple, au lieu du schéma de la figure 6 on obtient une ligne unidimensionnelle de symboles

$$q_1 \mid \alpha I q_2 q_2 \mid \beta I q_1 q_3 \parallel D q_1 \dots \qquad (\Omega)$$

Il est clair qu'à partir de cette ligne on peut toujours rétablir d'une façon bijective le tableau initial. De façon analogue, lors de l'examen des configurations on peut convenir de mettre un symbole d'état non pas sous la lettre examinée, mais immédiatement à sa gauche. La configuration IV de la figure 18 prend alors la forme

$$\mid\mid\mid q_4 \mid\mid .$$

Une telle représentation unidimensionnelle d'une configuration permet à volonté de rétablir d'une façon bijective sa forme initiale.

2. Les schémas fonctionnels et configurations ne dépendent aucunement de la forme des lettres figurant dans l'alphabet extérieur et dans celui des états. Ainsi, par exemple, si dans le tableau de la figure ou dans la ligne qui lui correspond on remplace partout la lettre β par la lettre b, cela ne modifie en aucune manière nos raisonnements. Il importe seulement de désigner des données différentes par des symboles différents et de pouvoir distinguer les lettres marquant les états de celles de l'alphabet extérieur.

On aurait bien sûr pu prendre des lettres autres que G, D, I pour désigner les décalages (à droite, à gauche, absence de décalage), mais on doit dire nettement quelle lettre désigne tel ou autre décalage. Cela s'explique par le fait que chacune des trois lettres symbolise une action bien déterminée que l'on ne peut remplacer par une autre.

Ceci dit, remplaçons dans la ligne Ω chaque lettre prise séparément par une suite composée de 0 et de 1 (*groupe codé*), en sorte que des lettres différentes soient remplacées par des groupes codés différents, mais que la même lettre, où qu'elle se trouve, soit partout remplacée par un seul et même groupe codé. Un tel remplacement aboutit à ce que la ligne Ω, par exemple, soit remplacée par Ω'. Pour pouvoir rétablir Ω à partir de Ω', le procédé de codage (correspondance fixée entre les lettres et les groupes codés) doit satisfaire aux conditions suivantes:

1) la ligne Ω' peut être divisée de façon bijective en groupes codés isolés;

2) on peut toujours savoir quels groupes codés se rapportent à chacune des lettres G, D, I prises séparément et distinguer les groupes codés choisis pour marquer les états de ceux de l'alphabet extérieur.

Le procédé de codage indiqué ci-dessous satisfait, il est clair, à ces deux conditons:

1. Comme groupes codés on prend $3 + k + m$ mots distincts de la forme

$$100\ldots01$$

(entre les unités il n'y a que des zéros).

Alors la division de la ligne Ω' en groupes codés se fait facilement et de manière bijective en choisissant des suites de zéros comprises entre deux unités.

2. La correspondance entre les groupes codés et les lettres initiales se fait d'après le tableau de codage suivant:

	Lettre	Groupe codé	
	G	101	
	I	1001	
	D	10001	
Alphabet extérieur	S_1	100001 4 zéros	Nombre pair de zéros supérieur à 2
	S_2	10000001 6 zéros	
	. . . .		
	S_k	10 . . . 01 $2(k+1)$ zéros	
Alphabet des états	q_1	1000001 5 zéros	Nombre impair de zéros, supérieur à 5
	q_2	100000001 7 zéros	
	. . . .		
	q_m	10 . . . 01 $2(m+1)+1$ zéros	

Dans un tel système de codage Ω' prend, dans le cas considéré, la forme:

|00000||0000||000000||00||0000000||0000000||0000||
||00000000||00||00000| . . .

Une telle ligne formée de 0 et 1 et établie pour un schéma fonctionnel ou des configurations s'appellera *chiffre d'un schéma fonctionnel* ou *chiffre de la configuration* respectivement. Un tel chiffre permet de rétablir aisément la forme initiale d'un schéma fonctionnel ou d'une configuration. Il est donc aisé de donner les schémas et configurations à l'aide de leurs chiffres. Il va de soi qu'à la place des 0 et des 1 on pourrait prendre deux autres signes quelconques, a et b, par exemple.

Il est maintenant facile de modifier les formulations des instructions 1-9 de la forme initiale de l'algorithme d'imitation indiqué ci-dessus pour obtenir un algorithme servant à transformer les chiffres du schéma de la machine imitée et le chiffre de la configuration initiale en le chiffre de la configuration qui en résulte. Bornons-nous à donner ici quelques illustrations:

Instruction 1. Examiner dans le chiffre de la configuration le groupe codé (unique) situé immédiatement à droite du groupe codé qui comporte un nombre impair de zéros.

Instructions 2 et 3. Chercher dans le chiffre du schéma un couple de groupes codés voisins identique au couple de groupes codés dans le chiffre de la configuration où est examiné le second groupe.

Instruction 6. Si dans le triplet de groupes codés considéré le deuxième est le groupe 1001, remplacer le groupe codé qui comporte un nombre impair de 0 dans le chiffre de la configuration par le troisième groupe codé du triplet considéré.

Une analyse plus poussée de cet algorithme permet de ramener chaque opération effectuée sur les groupes codés à une chaîne d'opérations standards réalisées sur une machine de Turing (remplacement d'un signe par un autre, décalage à un pas, etc.). Outre les signes 1 et 0 formant les chiffres, on fait également appel à d'autres lettres, pour séparer les chiffres les uns des autres, ou bien pour marquer provisoirement les 0 et les 1 (cf. algorithme d'Euclide), etc.

Un algorithme d'imitation ainsi détaillé s'avère finalement décrit par un schéma fonctionnel de Turing. C'est précisément le schéma de la machine universelle. Si une machine *A* résout un problème quelconque, la machine universelle est capable, elle aussi, de résoudre ce même problème à condition que l'on inscrive sur la bande, outre le chiffre des données initiales de ce problème, le chiffre du schéma de la machine *A*.

Ainsi, tous les schémas fonctionnels (ou leurs chiffres) admettent deux interprétations suivantes:

1) un schéma décrit le bloc logique d'une machine de Turing spéciale réalisant l'algorithme correspondant (c'est le point de vue que l'on a initialement adopté);

2) un schéma décrit le programme enregistré sur le ruban de la machine universelle pour réaliser l'algorithme correspondant.

Remarquons pour conclure que les machines à calculer électroniques modernes se construisent justement comme les machines universelles dans la mémoire desquelles on introduit, à côté des données initiales du problème posé, le programme de sa résolution.

La division de la mémoire en mémoires extérieure et intérieure est caractéristique des machines à calculer; de plus, la première est le plus souvent constituée en tambours et en bandes magnétiques servant à enregistrer l'information, tout comme on le fait pour l'enregistrement des sons au magnétophone. Mais à la différence de la machine de Turing dont la mémoire extérieure est illimitée (ruban illimité), celle des machines réelles est limitée (bande ou tambour magnétiques).

Il va de soi que cette différence de principe entre une machine à calculer réelle et une machine de Turing, qui représente une machine abstraite idéalisée, est insurmontable. Mais il importe de souligner que la mémoire extérieure d'une machine à calculer réelle peut être augmentée à l'infini sans changer la construction de la machine; pour cela il suffit de « coller » à la bande magnétique un « morceau » supplémentaire.

§ 11. PROBLÈMES INSOLUBLES PAR VOIE ALGORITHMIQUE

Le passage de la notion assez vague d'algorithme à la notion rigoureuse de machine de Turing, qui peut être donnée par son chiffre, permet aussi de préciser la question concernant la possibilité de résoudre par voie algorithmique (ou sur machine) certains problèmes. Cette question est à comprendre comme suit: est-ce qu'il existe ou non une machine de Turing capable de résoudre une classe donnée de problèmes? (pour l'expression « la machine de Turing résout une certaine classe de problèmes » cf. § 7).

Dans de nombreux cas la théorie des algorithmes donne

une réponse négative. L'un des premiers résultats de ce type obtenu en 1936 par le mathématicien américain Church se rapporte au problème de déductibilité en logique mathématique (cf. § 6).

● THÉORÈME DE CHURCH. *Le problème de déductibilité est insoluble par voie algorithmique.*

Ce théorème a non seulement permis de comprendre l'insuccès des tentatives entreprises antérieurement pour construire les algorithmes correspondants mais a même montré que ces tentatives n'auraient aucun sens.

Les démonstrations de l'insolubilité d'un problème, propres à la théorie des algorithmes, sont aussi rigoureuses que celles faites dans d'autres branches des mathématiques (l'impossibilité de réaliser la trisection d'un angle par le seul emploi de la règle et du compas, par exemple, ou bien la démonstration du fait que le côté d'un carré est incommensurable avec sa diagonale). A titre d'exemple nous allons esquisser ci-dessous la démonstration du *problème d'auto-application.*

Admettons que sur le ruban d'une machine de Turing soit enregistré son propre chiffre (c'est-à-dire le chiffre du schéma fonctionnel de la machine) introduit dans l'alphabet de la machine. Les deux cas suivants peuvent alors se présenter : 1) la machine est applicable à son propre chiffre, c'est-à-dire elle le transforme et, après avoir effectué un nombre fini de manipulations, s'arrête en émettant le signal d'arrêt ; 2) la machine est inapplicable à son chiffre, c'est-à-dire le signal d'arrêt n'apparaît jamais. D'où il s'ensuit que les machines (chiffres) elles-mêmes peuvent être divisées en deux classes : celles des machines de Turing auto-applicables et celles des machines non auto-applicables. Il se pose donc le problème général suivant.

● **Le problème d'auto-applicabilité.** *Etant donné un chiffre, déterminer si la machine qu'il détermine est auto-applicable ou non.*

C'est un problème typique de la construction d'un algorithme, car pour sa résolution il faut trouver une méthode générale (algorithme ou machine) permettant de savoir automatiquement d'après tout chiffre donné si cet algorithme est ou non auto-applicable.

● THÉORÈME. *Le problème d'auto-applicabilité est insoluble par voie algorithmique.*

Démonstration. Raisonnons par l'absurde. Supposons qu'une telle machine A existe. La machine A transforme alors tout chiffre auto-applicable en un symbole σ quelconque (représentant la réponse affirmative à la question de savoir si la machine A est auto-applicable ou non) et tout chiffre non auto-applicable en un autre symbole τ (représentant la réponse négative à cette même question). Dans ce cas on pourrait construire une autre machine B qui transforme toujours les chiffres non auto-applicables dans le symbole τ, alors que B n'est plus applicable aux chiffres auto-applicables. Pour cela on pouvait modifier le schéma de la machine B de façon qu'après l'apparition du symbole σ la machine au lieu d'émettre le signal d'arrêt répète indéfiniment ce même symbole.

Ainsi, B est applicable à tout chiffre non auto-applicable (en faisant toujours apparaître le symbole τ) mais ne l'est pas aux chiffres auto-applicables. Mais cela conduit à une contradiction. En effet: 1) supposons que la machine B soit auto-applicable, elle est alors applicable à son propre chiffre B qu'elle transforme en symbole τ; mais l'apparition de ce symbole signifie précisément que B est non auto-applicable; 2) supposons que B ne soit pas auto-applicable, alors elle n'est pas applicable à B, ce qui signifie que B (B') est auto-applicable. La contradiction ainsi obtenue démontre le théorème.

Les premiers résultats sur l'insolubilité algorithmique ont été obtenus pour des problèmes de logique mathématique (problème de déductibilité) et de théorie des algorithmes

(problème d'auto-applicabilité, par exemple). Néanmoins, il s'est avéré par la suite que de tels phénomènes ont également lieu lors de la résolution de problèmes apparemment plus restreints et qui apparaissent dans les branches les plus diverses des mathématiqus. Il faut avant tout signaler une série de problèmes algébriques se ramenant à diverses variantes du problème des mots qui ont été étudiées par des mathématiciens soviétiques.

Le problème d'équivalence des mots pour les calculs associatifs (cf. § 3) a été déjà formulé en 1914 par le mathématicien norvégien Tuet qui a également proposé l'algorithme pour déterminer l'équivalence des mots dans certains calculs associatifs particuliers. Dès lors on a entrepris bien des tentatives pour construire un algorithme général qui permettrait d'établir l'équivalence ou la non-équivalence de n'importe quel couple de mots dans tout calcul associatif. En 1946 et 1947, le mathématicien soviétique A. Markov et, indépendamment de lui, le mathématicien américain E. Post ont construit les exemples concrets des calculs associatifs pour lesquels la résolution algorithmique du problème d'équivalence des mots s'est avérée impossible. A plus forte raison il n'existe pas d'algorithme servant à déterminer l'équivalence des mots dans tout calcul associatif. A partir de ce résultat A. Markov et ses disciples ont démontré l'impossibilité de construire de tels algorithmes pour une classe plus large des calculs associatifs.

En 1955, le monde mathématique a été impressionné par le résultat de P. Novikov concernant l'insolubilité algorithmique du problème de l'identité de la théorie des groupes *). Du point de vue formel ce problème représente un cas particulier du problème de l'équivalence des mots dans un calcul assocaitif *). On ne considère notamment que les calculs associatifs où chaque lettre *a* de l'alphabet admet, parmi toutes les substitutions permises, une substitution

---

*) Ce travail de *P. Novikov* lui a valu le prix Lénine de 1957.

de la forme

$$a\alpha - \Lambda,$$

où α est une lettre quelconque de ce même alphabet (coïncidant éventuellement avec a).

Le sens de cette restriction devient clair si l'on considère, par analogie avec l'exemple 4 du § 3, les mots de tout calcul associatif comme des transformations complexes qui s'obtiennent par multiplication de transformations élémentaires données par des lettres correspondantes qui forment le mot considéré. Dans ce cas le mot vide Λ désigne la transformation identique qui ne modifie rien (cf. § 3) et la présence de substitutions permises de la forme $a\alpha - \Lambda$ signifie qu'à chaque transformation élémentaire (donnée par la lettre a) correspond une transformation élémentaire (donnée par la lettre α) telle que l'application successive de ces deux transformations donne la transformation identique. Sans entrer dans les détails, bornons-nous à noter que l'étude de tels ensembles de transformations, appelés groupes de transformations, sont d'un grand intérêt théorique et pratique et la notion de groupe est l'une des notions fondamentales des mathématiques modernes.

Il est tout de même à noter que le résultat très important de Markov-Post ne permet pas en lui-même de tirer des conclusions sur le problème de l'identité de la théorie des groupes. La cause en est que les calculs associatifs particuliers, pour lesquels A. Markov et E. Post ont démontré l'insolubilité algorithmique du problème d'équivalence, ne satisfont pas à la restriction imposée plus haut qui joue un rôle fondamental pour le problème de l'identité dans la théorie

---

*) Cela signifie que s'il existait un algorithme servant à démontrer l'identité des mots dans un calcul associatif, celui-ci permettrait également d'établir l'identité des mots dans un groupe. Cependant, l'insolubilité algorithmique du problème de l'identité des mots dans un calcul associatif n'entraîne aucunement celle du problème correspondant de la théorie des groupes. (*Note du rédacteur*).

des groupes. Les résultats de Markov-Post n'excluent donc pas qu'un tel algorithme puisse exister. Ainsi, on avait poursuivi les recherches dans l'espoir de trouver un tel algorithme, quand P. Novikov a fait connaître les résultats de ses travaux. Il a construit un exemple particulier d'un calcul associatif, vérifiant la restriction mentionnée plus haut, mais dont l'équivalence des mots est insoluble par voie algorithmique ; il est à plus forte raison impossible d'avoir un algorithme unique pour tous les groupes considérés.

Les exemples construits par A. Markov et P. Novikov pour réfuter la possibilité de résoudre les problèmes considérés par voie algorithmique sont très longs et admettent des centaines de substitutions admissibles. Le problème s'est posé de trouver des exemples plus simples. Tout récemment ce problème a été brillamment résolu par le jeune mathématicien de Léningrad G. Tséitine qui a trouvé un calcul (mentionné dans le § 3) à sept substitutions permises et dont le problème de l'équivalence des mots n'admet pas de solution algorithmique.

L'existence des problèmes insolubles par voie algorithmique a mis la science dans une situation telle que si un mathématicien cherche à construire un algorithme, il doit compter avec le fait qu'un tel algorithme peut ne pas exister. Ainsi, la recherche d'un algorithme va de pair avec les efforts déployés pour démontrer l'absence d'un tel algorithme. Les progrès réalisés dans l'une de ces voies permettront donc de trouver un algorithme, ou bien de démontrer l'insolubilité algorithmique du problème considéré.

Au § 1 nous avons formulé le problème de Hilbert sur les équations diophantiennes. Les efforts pour construire cet algorithme, tous d'ailleurs vains, se sont poursuivis au cours d'un demi-siècle. On a également entrepris récemment de démontrer l'absence d'un tel algorithme et, bien qu'on n'ait pas obtenu de résultat définitif, il est quand même probable que le problème de Hilbert soit aussi insoluble par voie algorithmique.

CONCLUSION

Faisons pour terminer quelques remarques d'ordre général.

1. Premièrement, les théorèmes sur l'insolubilité algorithmique d'une classe ou d'une autre de problèmes ne doivent pas faire tomber dans l'agnosticisme. En effet, chacun de ces théorèmes concerne toute une classe de problèmes dont l'insolubilité est établie par une méthode effective unique, par un algorithme.

Cela ne signifie pas que parmi les problèmes isolés réunis dans cette classe il y a ceux qui ne sont pas résolubles. Ainsi, par exemple, du théorème démontré plus haut il ne faut pas déduire qu'il existe un tel chiffre dont on ne puisse en principe pas dire s'il est ou non auto-applicable.

Cela signifie seulement que la classe de problèmes considérée est à tel point générale et large qu'il n'existe tout simplement pas d'algorithme unique capable de résoudre tous les problèmes d'un type donné. Dans ce cas l'objectif des recherches mathématiques se ramène à la création successive d'algorithmes plus universels permettant de ramener au calcul automatique des sous-classes toujours plus larges des problèmes du type donné.

2. Deuxièmement, ces théorèmes montrent que les mathématiques ne se ramènent pas à la seule recherche des algorithmes et que le processus de la connaissance mathématique ne peut pas être entièrement automatisé. Aujourd'hui même de nombreux problèmes de certaines branches spécialisées des mathématiques (théorie des groupes à un nombre fini de génératrices par exemple, etc.) ne peuvent être résolus par aucune automate (c'est-à-dire par aucune machine de Turing à un nombre fini d'états et à mémoire limitée). Ainsi, toutes les affirmations sur la possibilité de remplacer le génie créateur des savants par une machine sont, il est clair, absurdes.

3. Il faut tout de même reconnaître que le domaine d'application des processus algorithmiques est assez large et dépasse le cadre des processus calculatoires rencontrés en mathématiques. De plus, pour bien des processus, que l'on considère généralement comme extrêmement difficiles et compliqués, on peut, du point de vue théorique, construire des algorithmes fondés sur une idée assez simple ; les difficultés pratiques auxquelles on se heurte lors de la réalisation de ces processus sont dues au fait que de tels algorithmes s'avèrent trop longs et nécessitent un nombre fort important d'opérations à effectuer (bien que très élémentaires). Cette dernière remarque se rapporte en particulier aux jeux (jeu d'échecs, par exemple), où le gain est lié à l'habileté d'examiner le plus grand nombre de variantes possibles pour en choisir une optimale.

L'apparition des machines à calculer rapides a permis d'augmenter considérablement le nombre d'algorithmes réalisables en pratique.

4. Avant de terminer notre petit exposé soulignons une fois de plus que chaque machine à calculer réelle ne représente qu'un modèle approché de la machine de Turing. Plus précisément, toutes les machines réelles sont à mémoire extérieure limitée, tandis que dans les schémas de la machine de Turing il s'agit d'une bande infinie. Il va de soi que la réalisation technique d'une mémoire illimitée est impossible mais une augmentation considérable des mémoires des machines à calculer est non seulement souhaitable mais, de plus, possible. C'est précisément dans le sens d'une augmentation de la mémoire extérieure et de la vitesse de calcul que les automates modernes ont certainement les plus belles perspectives.

E. Ventsel

ÉLÉMENTS DE LA THÉORIE DES JEUX

§ 1. OBJET DE LA THÉORIE DES JEUX. NOTIONS FONDAMENTALES

Certains problèmes pratiques (dans les domaines économique, militaire, etc.) conduisent à analyser des situations qui opposent deux (ou même plus) parties poursuivant des buts antagoniques et le résultat de toute action de l'une de ces parties dépend du comportement de l'autre. Ces situations seront désignées dans la suite par « situations de conflits ».

Les exemples de situations de conflits abondent dans les domaines les plus divers de la vie courante. Toute situation née d'opérations militaires est évidemment une situation de conflits: chaque partie fait tout son possible pour empêcher un succès adverse. De même, les situations liées au choix d'un système d'armements, aux modes de leurs applications militaires et, d'une façon générale, à l'élaboration des opérations stratégiques sont également des situations de conflits: chaque décision doit être prise dans l'éventualité d'une réponse adverse la moins avantageuse pour nous. Certaines situations économiques (surtout en présence d'une libre concurrence) font partie des situations de conflits ; les adversaires étant ici les firmes commerciales, industrielles, etc.

Le besoin d'analyser ces situations a donné corps à tout un appareil mathématique particulier. La théorie des jeux n'est au fond rien d'autre que la théorie mathématique des situations de conflits dont l'objet est d'élaborer des recommandations rationnelles sur la manière d'agir de chaque partie impliquée dans une situation de conflits.

Chaque situation de conflits réelle est particulièrement compliquée, car toute une série de facteurs accessoires rendent son analyse bien difficile. Ainsi, pour que ces situations se prêtent à une analyse mathématique, il importe de faire abstraction de tous les facteurs secondaires de manière à en réaliser un modèle simplifié, formalisé. Un tel modèle s'appellera « jeu ».

Un jeu se distingue d'une situation de conflits réelle par le fait qu'il obéit à des *règles* bien définies. Les modèles formalisant les situations de conflits et qui constituent les *jeux* proprement dits, sont connus des hommes depuis longtemps : citons les jeux d'échecs, de dames, de cartes, etc. Tous ces jeux revêtent la forme d'une compétition, soumise à des règles bien déterminées, qui se termine par la « victoire » (gain) d'un joueur.

Ces jeux formellement réglementés et spécialement créés sont particulièrement aptes à illustrer et faire assimiler les notions fondamentales de la théorie des jeux. La terminologie qui leur sera empruntée est également valable pour analyser d'autres situations de conflits : les participants d'un jeu sont conventionnellement appelés « joueurs » et le résultat d'une opposition des intérêts — « gain » d'un des joueurs.

Un jeu peut opposer les intérêts de deux ou de plusieurs adversaires : le premier cas est celui du jeu à deux participants, le deuxième, à plusieurs participants. Les participants d'un jeu à plusieurs personnes peuvent former des alliances permanentes ou temporaires. La formation de deux alliances permanentes transforme un jeu à plusieurs participants en jeu à deux. Les jeux à deux étant les plus importants, nous n'examinerons qu'eux dans ce livre.

Commençons l'étude de la théorie élémentaire des jeux par l'exposé des notions fondamentales. Nous examinerons un jeu à deux personnes A et B dont les intérêts sont opposés. Par « jeu » nous entendrons une série d'actions des joueurs A et B. Pour rendre un jeu mathématiquement analysable il faut formuler d'une façon précise les *règles du jeu*. Les « règles du jeu » sont un système de conditions réglementant le comportement de chacun des deux joueurs, le volume de l'information que possède chaque joueur sur la conduite de l'autre, l'alternance des « coups » (solutions isolées que l'on adopte au cours d'un jeu), ainsi que le *résultat* ou *issue* auquel aboutit l'ensemble considéré des

coups. Bien que ce résultat (gain ou perte) ne s'exprime pas toujours quantitativement, on peut, néanmoins, en adoptant une échelle numérique, le traduire par un nombre quelconque. Ainsi, dans le jeu d'échecs on attribue par convention +1 à une partie gagnée, —1 à une partie perdue et 0 à une partie nulle.

Un jeu est dit *à somme nulle* si chaque adversaire gagne ce que l'autre perd, c'est-à-dire, si la somme des gains des deux partenaires est nulle. Les intérêts des joueurs dans un jeu à somme nulle sont diamétralement opposés. Seuls ces jeux feront l'objet de ce livre.

Vu que dans un jeu à somme nulle le gain d'un joueur est égal à celui de l'autre affecté du signe opposé, on peut donc en analysant un tel jeu se borner à examiner le gain de l'un des deux joueurs, disons celui de A. Pour plus de commodité, nous conviendrons dans la suite d'appeler le joueur A « nous » et le joueur B « adversaire ».

Nous supposerons également que le joueur A (« nous ») est toujours « gagnant » et le joueur B (« adversaire ») toujours « perdant ». Cette convention purement formelle d'ailleurs ne donne, en vérité, aucun avantage réel au premier joueur ; il est aisé de voir qu'un changement de signe opposé inverse la situation.

Nous allons présenter le déroulement du jeu comme une succession d'étapes ou de « coups ». Dans la théorie des jeux on appelle *coup* le choix de l'une des manières possibles d'agir, conformes aux règles du jeu. Les coups se divisent en *personnels* et *fortuits*.

Le *coup personnel* est le choix conscient et la réalisation d'un des coups à la disposition d'un joueur dans une situation déterminée. Toute marche d'une pièce d'échecs est un coup personnel. Le joueur d'échecs qui a le trait fait consciemment choix d'une possibilité de joueur que lui offre la disposition des pièces. Ces possibilités sont déterminées par les règles du jeu et dépendent de l'ensemble des coups déjà effectués par les deux adversaires.

Le *coup fortuit* est le choix d'un coup possible réalisé non pas par le joueur, mais grâce à un mécanisme quelconque de choix arbitraire (jeu de « pile ou face », « de dés », le battage et la donne des cartes, etc). Ainsi, la donne de la première carte dans une partie de poker est un coup fortuit sur 32 possibilités équivalentes.

Pour qu'un jeu soit mathématiquement défini il faut que ses règles déterminent pour chaque coup fortuit la *répartition de probabilités* des chances possibles.

Certains jeux sont composés uniquement de coups fortuits (on les appelle jeux de hasard pur) ou uniquement de coups personnels (les jeux d'échecs, de dames). La plupart des jeux de cartes sont des jeux mixtes, c'est-à-dire comportent des coups fortuits et des coups personnels.

Les jeux peuvent être classés non seulement d'après le type des coups (personnels, fortuits) mais également selon le caractère et le volume de l'information dont dispose chaque joueur sur la manière d'agir de son partenaire. On appelle *jeux à information parfaite* les jeux où chaque joueur est parfaitement au courant de tous les coups précédents aussi bien personnels que fortuits : les jeux d'échecs, de dames, etc.

La plupart des jeux présentant un intérêt pratique n'appartiennent pas à la classe des jeux à information parfaite, étant donné que l'ignorance du comportement de l'adversaire est un élément essentiel dans les situations de conflits.

L'une des notions fondamentales de la théorie des jeux est la notion de « stratégie ».

La *stratégie* est l'ensemble des règles impliquant au joueur ayant le trait un coup personnel unique en fonction de la tournure du jeu.

Etudions plus en détail cette notion de stratégie.

Ordinairement, le choix d'un coup personnel (ou décision) répond à une situation concrète. Or, théoriquement rien ne change si l'on s'imagine qu'un joueur prend toutes ses décisions *par avance*. Pour cela le joueur aurait dû dresser préalablement la liste de toutes les situations possibles et

prendre les décisions adéquates. En principe (voire même en pratique) cela est parfaitement possible pour tout jeu. Si un tel système est adopté on dit qu'un joueur a choisi une certaine *stratégie*. Le joueur qui a fait choix d'une stratégie peut ne pas participer personnellement au jeu, et se faire remplacer par une tierce personne (l'arbitre par exemple) qui se conformera à la liste des règles élaborées par ce joueur. Parfois on confiera cette mission à une calculatrice qu'on aura munie d'un programme approprié. C'est ainsi qu'on voit aujourd'hui les calculatrices jouer aux échecs.

La notion de « stratégie » n'a de sens que si le jeu comporte des coups personnels; pour les jeux de hasard pur cette notion n'a plus de raison d'être.

Les jeux se divisent en « finis » et « infinis » selon le nombre de stratégies possibles.

Un jeu dans lequel chaque adversaire dispose d'un nombre fini de stratégies est appelé *fini*.

Un jeu fini dans lequel le joueur A dispose de m stratégies et le joueur B de n stratégies est appelé jeu de $m \times n$.

Examinons donc un jeu de $m \times n$ opposant A et B (« nous » et l'« adversaire »).

Désignons nos stratégies par $A_1, A_2, \ldots, A_m$; celles de l'adversaire par $B_1, B_2, \ldots, B_n$.

Admettons que chaque joueur ait déjà choisi sa propre stratégie: A_i (pour « nous ») et B_j (pour l'« adversaire »).

Si le jeu ne comporte que des coups personnels, le choix des stratégies A_i, B_j en détermine le résultat d'une façon univoque — nous sommes donc gagnants. Désignons notre gain par a_{ij}.

Si le jeu comporte des coups personnels et des coups fortuits, le gain pour un couple de stratégies A_i, B_j est une valeur aléatoire dépendant de l'issue de tous les coups fortuits. Dans ce cas on estime un gain par sa *valeur moyenne* (espérance mathématique). Nous allons désigner par le même signe a_{ij} le gain lui-même (pour un jeu sans coups fortuits) et sa valeur moyenne (pour un jeu à coups fortuits).

Supposons connues les valeurs a_{ij} du gain (ou de sa valeur moyenne) pour chaque couple de stratégies. Ces valeurs peuvent être inscrites dans un tableau rectangulaire (matrice) dont les lignes correspondent à nos stratégies (A_i) et les colonnes à celles de l'adversaire (B_j). Un tel tableau s'appelle *matrice de paiement* ou *matrice du jeu* tout simplement.

La matrice du jeu de $m \times n$ s'écrit donc:

A \ B	B_1	B_2		B_n
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
⋮	⋮	⋮	⋯	⋮
A_m	a_{m1}	a_{m2}	...	a_{mn}

Nous désignerons la matrice d'un jeu par $\| a_{ij} \|$.

Considérons maintenant quelques exemples fort élémentaires de jeux.

● EXEMPLE 1. Deux joueurs A et B retournent à leur gré, sans se regarder d'ailleurs, chacun sa pièce. S'il y a concordance le joueur A prend les deux pièces; sinon c'est le joueur B qui les prend. On demande d'analyser ce jeu et de composer sa matrice.

● SOLUTION. Ce jeu ne se compose que de deux coups: le nôtre et celui de l'adversaire, tous les deux étant personnels. Il n'appartient pas aux jeux à information parfaite, car le joueur qui a le trait ne connaît rien du coup réalisé par son rival.

Étant donné que chaque participant ne dispose que d'un seul coup personnel, la stratégie d'un joueur se ramène donc au choix de ce coup unique.

Nous avons deux stratégies: A_1, choisir pile et A_2, choisir face; l'adversaire dispose aussi des mêmes stratégies: B_1 — pile et B_2 — face. Ainsi, le jeu considéré est un jeu de 2×2. Désignons le gain d'une pièce par $+1$. Le tableau de ce jeu se représente ainsi:

A \ B	B_1 (p)	B_2 (f)
A_1 (p)	1	—1
A_2 (f)	—1	1

Ce jeu, tout élémentaire qu'il soit, permet d'assimiler certaines notions importantes de la théorie des jeux.

Supposons d'abord que les concurrents ne jouent qu'une seule fois. Il est alors absurde de dire que certaines « stratégies » sont plus raisonnables que d'autres. Chaque joueur peut prendre une décision quelconque avec les mêmes chances de succès. Mais il n'en est pas de même si l'on continue à jouer.

En effet, admettons que nous (joueur A) ayons adopté une certaine stratégie (A_1, par exemple) que nous suivrons toujours. Dans ce cas dès les premiers coups, l'adversaire devinera certainement notre stratégie et ne manquera pas de nous placer dans une situation désavantageuse, c'est-à-dire qu'il choisira face. Nous n'avons donc aucune raison d'adopter toujours la même stratégie; pour ne pas être en perte il nous faut choisir tantôt pile tantôt face. Mais si nous alternons régulièrement les piles et les faces (exemple, une fois pile, une fois face), l'adversaire peut également deviner cette stratégie et choisir la meilleure parade. Il est évident

que le procédé le plus sûr pour que l'adversaire ne puisse pas deviner notre stratégie est celui où nous ne savons pas nous-mêmes notre coup suivant (ce qu'on peut faire, par exemple, en jetant notre pièce en l'air). Ainsi, ces raisonnements purement intuitifs nous amènent à l'une des notions les plus fondamentales de la théorie des jeux, celle de « stratégie mixte », c'est-à-dire une stratégie faisant alterner les stratégies « pures » — ici A_1 et A_2 — au hasard et avec une certaine fréquence. Dans l'exemple étudié on comprend aisément que pour des considérations de symétrie les stratégies A_1 et A_2 doivent alterner avec la même fréquence ; dans des jeux plus compliqués la solution est loin d'être aussi triviale.

● EXEMPLE 2. Deux joueurs A et B notent simultanément et indépendamment l'un de l'autre un des trois chiffres: 1, 2 ou 3.

Si la somme des chiffres notés est paire, B la paye à A en francs, si elle est impaire alors A la paye à B. On demande d'analyser ce jeu et de composer sa matrice.

● SOLUTION. Ce jeu se compose de deux coups personnels. Nous avons trois stratégies que nous noterons A_1, A_2 et A_3 selon que le chiffre choisi est 1, 2 ou 3. L'adversaire possède ces mêmes stratégies. Nous sommes donc en présence d'un jeu de 3 × 3 dont la matrice est donnée ci-dessous.

A \ B	B_1	B_2	B_3
A_1	2	—3	4
A_2	—3	4	—5
A_3	4	—5	6

Il est évident que dans ce cas aussi en répondant à l'une quelconque de nos stratégies, l'adversaire cherchera à nous mettre dans la situation la plus désavantageuse. En effet, si nous choisissons, par exemple, la stratégie A_1, l'adversaire nous opposera la stratégie B_2 ; à la stratégie A_2 il répliquera par B_3 ; à la stratégie A_3 il répondra par B_2 ; donc, quelle que soit la stratégie choisie nous serons inévitablement perdants *). La résolution de ce jeu (c'est-à-dire le choix des stratégies les plus avantageuses de deux joueurs) sera donnée au § 5.

● EXEMPLE 3. Nous possédons trois types d'armes: A_1, A_2, A_3 ; l'adversaire possède trois types d'avions: B_1, B_2, B_3. Notre but est d'atteindre l'avion, celui de l'adversaire, d'éviter d'être abattu. Les probabilités d'atteindre les avions B_1, B_2, B_3 avec l'arme A_1 sont respectivement 0,9 ; 0,4 et 0,2 ; avec l'arme A_2 — 0,3 ; 0,6 et 0,8 ; avec l'arme A_3 — 0,5 ; 0,7 et 0,2. On demande de formuler ces données en termes de théorie des jeux.

● SOLUTION. Cette situation peut être considérée comme un jeu de 3×3 avec deux coups personnels et un coup aléatoire. Notre coup personnel c'est le choix de l'arme ; le coup personnel de l'adversaire c'est le choix de l'avion. Le coup aléatoire c'est l'emploi de l'arme qui pourra ou non abattre l'avion. Notre gain sera égal à l'unité si l'avion est abattu et à zéro dans le cas contraire. Nos stratégies sont les trois types d'armes ; les stratégies de l'adversaire sont les trois types d'avions. L'espérance mathématique du gain pour chaque couple de stratégies données n'est rien d'autre que la probabilité d'abattre un avion donné avec l'arme choisie. La matrice de ce jeu est donnée ci-dessous.

---

*) Il faut dire que notre adversaire se trouve dans une situation aussi critique.

A \ B	B_1	B_2	B_3
A_1	0,9	0,4	0,2
A_2	0,3	0,6	0,8
A_3	0,5	0,7	0,2

Le but de la théorie des jeux est d'élaborer des recommandations sur le comportement des joueurs dans les situations de conflits, c'est-à-dire la détermination de la « stratégie optimale » pour chaque joueur.

On appelle *stratégie optimale* celle qui garantit à un joueur donné, lorsque le jeu se répète, un gain moyen maximal (ou, ce qui revient au même, une perte moyenne minimale). Le choix d'une telle stratégie est basé sur l'hypothèse que notre adversaire est pour le moins aussi raisonnable que nous et fera tout pour nous empêcher de réaliser nos objectifs.

Dans la théorie des jeux toutes les recommandations sont élaborées en partant de ces principes ; on ne tiendra donc pas compte du facteur risque inhérent à toute stratégie réelle, de même que des fautes et erreurs de calcul de chaque joueur.

Comme tout modèle mathématique d'un phénomène complexe la théorie des jeux a ses propres restrictions, dont la principale est que le gain est artificiellement réduit à un unique nombre. Mais dans la plupart des situations de conflits réelles l'élaboration d'une stratégie raisonnable nécessite la prise en compte de plusieurs paramètres numériques, facteurs du succès d'une entreprise. Une stratégie optimale pour certains critères ne l'est pas forcément pour d'autres. Néanmoins, tout en se rendant compte de ces restrictions

pour ne pas se conformer aveuglément aux recommandations élaborées par les méthodes de la théorie des jeux, on peut tout de même utiliser d'une manière raisonnable son appareil mathématique pour former une stratégie « admissible » sinon précisément « optimale ».

§ 2. VALEURS INFÉRIEURE ET SUPÉRIEURE D'UN JEU. PRINCIPE MINIMAX

Considérons un jeu de $m \times n$ dont la matrice a la forme

A \ B	B_1	B_2		B_n
A_1	a_{11}	a_{12}		a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
⋮		...	...	
A_m	a_{m1}	a_{m2}	...	a_{mn}

Désignons par i le numéro de notre stratégie, par j celui de la stratégie de notre adversaire.

Le problème consiste à définir notre stratégie optimale. Pour cela analysons successivement chacune de nos stratégies en partant de A_1. En choisissant la stratégie A_i il nous faut toujours tenir compte que notre adversaire lui opposera celle des stratégies B_j qui minimisera notre gain a_{ij}. Définissons donc ce gain, c'est-à-dire le plus petit des nombres

a_{ij} de la i-ième ligne. Désignons-le par α_i:

$$\alpha_i = \min_j a_{ij}. \tag{2.1}$$

Le symbole $\min\limits_j$ désigne ici le minimum des valeurs du paramètre donné pour tous les j possibles.

Ecrivons les nombres α_i, à droite de notre matrice, en colonne:

$A \backslash B$	B_1	B_2	...	B_n	α_i
A_1	a_{11}	a_{12}	...	a_{1n}	α_1
A_2	a_{21}	a_{22}	...	a_{2n}	α_2
⋮	⋮	⋮	⋮	⋮	⋮
A_m	a_{m1}	a_{m2}	...	a_{mn}	α_m
β_j	β_1	β_2	...	β_m	

En adoptant une stratégie quelconque A_i, nous devons espérer que, compte tenu des actions raisonnables de l'adversaire, notre gain ne dépassera pas α_i. Il est naturel qu'en agissant de la façon la plus prudente et en supposant que notre adversaire sera raisonnable (c'est-à-dire en évitant le moindre risque) nous devons opter pour la stratégie A_i telle qui maximise le nombre α_i. Désignons cette valeur maximale de α par

$$\alpha = \max_i \alpha_i$$

ou, si l'on tient compte de (2.1),

$$\alpha = \max_i \min_j \alpha_{ij}.$$

Le nombre α est appelé *valeur inférieure du jeu* ou *gain maximin* ou, enfin, *maximin* tout court.

Le nombre α se trouve sur une ligne déterminée de la matrice: la stratégie du joueur A qui correspond à cette ligne est appelée *stratégie maximin.*

Il est évident qu'en choisissant la stratégie maximin *nous nous assurons dans tous les cas*, quelle que soit la stratégie de l'adversaire, *un gain dont la valeur n'est pas inférieure à* α. C'est pour cette raison d'ailleurs que le nombre α est appelé « valeur inférieure du jeu ». En d'autres termes, c'est le gain minimum garanti que l'on peut s'assurer en adoptant la stratégie la plus prudente.

Il est évident que les mêmes raisonnements valent pour l'adversaire B. Etant donné que l'adversaire cherche à minimiser notre gain, il ne doit donc concevoir que des stratégies qui lui procureront un gain maximal. Inscrivons donc les valeurs maximales de a_{ij} au bas de chaque colonne:

$$\beta_j = \max_i a_{ij}$$

et déterminons la valeur minimale de β_j:

$$\beta = \min_j \beta_j$$

ou

$$\beta = \min_j \max_i a_{ij}.$$

Le nombre β s'appelle *valeur supérieure du jeu*, ou *minimax*. La stratégie de l'adversaire qui correspond au gain minimax est appelée *stratégie minimax.*

En adoptant sa stratégie minimax la plus prudente, l'adversaire *se garantit à coup sûr*, quelle que soit notre stratégie, *une perte dont la valeur ne dépassera pas* β.

Dans la théorie des jeux et ses applications, le principe de prudence qui dicte aux joueurs le choix des stratégies correspondantes (maximin et minimax) est souvent appelé « principe minimax ». Les stratégies maximin et minimax sont parfois désignées par un terme commun « stratégies minimax ».

Définissons à titre d'exemple les valeurs inférieure et supérieure du jeu et les stratégies minimax pour les exemples 1, 2 et 3 du § 1.

● EXEMPLE 1. Dans l'exemple 1 du § 1, nous avons examiné le jeu dont la matrice est

$A \backslash B$	B_1	B_2	α_i
A_1	1	—1	—1
A_2	—1	1	—1
β_j	1	1	

Etant donné que les grandeurs α_i et β_j sont constantes et respectivement égales à —1 et +1, les valeurs inférieure et supérieure du jeu sont de même égales à —1 et +1 :

$$\alpha = -1; \quad \beta = +1.$$

Toute stratégie du joueur A est une stratégie maximin et toute stratégie du joueur B est une stratégie minimax. La conclusion est triviale : en choisissant l'une quelconque de leurs stratégies le joueur A de même que B se garantissent une perte ne dépassant pas 1.

● EXEMPLE 2. La matrice du jeu de l'exemple 2 du § 1 est :

A \ B	B_1	B_2	B_3	α_i
A_1	2	—3	4	—3
A_2	—3	4	—5	—5
A_3	4	—5	6	—5
β_j	4	4	6	

La valeur inférieure du jeu est $\alpha = -3$, la valeur supérieure $\beta = 4$. Notre stratégie maximin est A_1 ; en l'adoptant systématiquement nous pouvons donc nous assurer un gain au moins égal à —3 (une perte maximale égale à 3). La stratégie minimax de l'adversaire est l'une quelconque des stratégies B_1 et B_2 ; en les adoptant d'une façon systématique il peut s'assurer dans tous les cas une perte ne dépassant pas 4. Si nous refusons de jouer notre stratégie maximin (si l'on opte pour la stratégie A_2, par exemple), l'adversaire en répliquant aussitôt par la stratégie B_3 peut nous infliger une sanction, en réduisant notre gain à —5 ; de même, si l'adversaire refuse sa stratégie minimax il portera sa perte à 6.

● EXEMPLE 3. La matrice du jeu examiné dans l'exemple 3 du § 1 est :

Les valeurs inférieure et supérieure du jeu sont respectivement $\alpha = 0{,}3$ et $\beta = 0{,}7$. La plus prudente de nos stratégies (stratégie maximin) est A_2 ; en utilisant l'arme A_2 nous

B / A	B_1	B_2	B_3	α_i
A_1	0,9	0,4	0,2	0,2
A_2	0,3	0,6	0,8	0,3
A_3	0,5	0,7	0,2	0,2
β_j	0,9	0,7	0,8	

sommes certains d'abattre l'avion en moyenne dans 0,3 des cas. La stratégie la plus prudente (stratégie minimax) de l'adversaire est B_2 ; en choisissant ce type d'avion l'adversaire est assuré de n'être abattu que dans 0,7 des cas.

Ce dernier exemple permet d'illustrer l'une des propriétés fondamentales des stratégies minimax : leur instabilité.

Admettons que nous avons choisi la stratégie maximin A_2 et l'adversaire sa stratégie minimax B_2. Tant que les deux participants joueront ces stratégies, le gain moyen restera égal à 0,6 ; il dépasse la valeur inférieure mais n'atteint pas la valeur supérieure. Mais si l'adversaire apprend notre stratégie A_2, il lui opposera aussitôt la stratégie B_1, réduisant notre gain à 0,3. Mais nous pouvons opposer la stratégie A_1 à B_1, ce qui nous assurera le gain 0,9, etc.

Si donc les deux participants adoptent leurs stratégies minimax, la situation devient instable et l'équilibre peut être troublé lorsqu'on apprend la stratégie de l'adversaire.

Néanmoins il existe des jeux à stratégies minimax stables. Ce sont des jeux dans lesquels la valeur inférieure est

égale à la valeur supérieure:

$$\alpha = \beta.$$

Si donc la valeur inférieure d'un jeu est égale à celle supérieure, cette grandeur générale est alors appelée *valeur nette du jeu* (ou, tout simplement, *valeur du jeu*) ; nous la désignerons dans la suite par ν.

Examinons maintenant un exemple. Soit un jeu de 4×4 donné par la matrice:

A \ B	B_1	B_2	B_3	B_4	α_i
A_1	0,4	0,5	0,9	0,3	0,3
A_2	0,8	0,4	0,3	0,7	0,3
A_3	0,7	0,6	0,8	0,9	0,6
A_4	0,7	0,2	0,4	0,6	0,2
β_j	0,8	0,6	0,8	0,9	

Déterminons la valeur inférieure:

$$\alpha = 0{,}6\,;$$

et celle supérieure

$$\beta = 0{,}6.$$

Celles-ci étant égales, ce jeu admet une valeur nette, dont la grandeur est $\alpha = \beta = \nu = 0{,}6$.

L'élément 0,6 est à la fois *minimum dans sa ligne et maximum dans sa colonne*. En géométrie, on appelle point selle

d'une surface le point jouissant de la propriété analogue, c'est-à-dire dont une coordonnée est minimale et l'autre maximale. L'élément d'une matrice de jeu jouissant de cette propriété s'appelle point selle ou *point d'équilibre de la matrice* et le jeu est alors appelé *jeu équilibré*.

Au point selle correspond un couple de stratégies minimax (dans l'exemple considéré ce sont A_3 et B_2). Ces stratégies sont appelées *optimales* et leur ensemble *solution de jeu*.

La solution de jeu possède la propriété remarquable suivante. Si l'un des joueurs (A par exemple) s'en tient toujours à sa stratégie optimale et l'autre (B) s'écarte, d'une façon ou d'une autre, de la sienne, alors le *joueur B se place dans une situation qui ne sera jamais à son avantage*; tout écart du joueur B de sa stratégie optimale peut, dans le meilleur cas, assurer le même gain à l'adversaire et, au pis, l'augmenter.

Si, par contre, le joueur B s'en tient constamment à sa stratégie optimale, et le joueur A s'écarte de sa stratégie optimale, la situation tourne à l'avantage de B.

Cette affirmation est bien facile à vérifier sur l'exemple du jeu équilibré considéré.

Comme on l'a vu plus haut les stratégies minimax des jeux équilibrés sont dans un sens « stables »: si l'un des participants s'en tient à sa stratégie minimax, tout écart de l'autre joueur de sa stratégie mimimax est à son désavantage. Il est à remarquer que même si l'un des joueurs sait bien que l'adversaire a choisi sa stratégie optimale, il ne peut rien changer dans sa propre conduite: s'il ne veut pas nuire à ses propres intérêts, il doit toujours jouer sa stratégie optimale. Un couple de stratégies optimales définit, dans un jeu équilibré, une sorte de « position d'équilibre »: si un joueur renonce à sa stratégie optimale il se met dans une situation qui le contraindra à y revenir.

Ainsi, tout jeu équilibré admet une solution déterminant un couple de stratégies optimales (une par joueur) dont voici les propriétés:

1) Si les deux participants s'en tiennent à leurs stratégies optimales, le gain moyen est égal à la valeur du jeu ν.

2) Si l'un des joueurs s'en tient à sa stratégie optimale et l'autre renonce à la sienne, ce dernier est irrémédiablement perdant.

Les jeux équilibrés présentent un intérêt aussi théorique que pratique.

Dans la théorie des jeux, on démontre, en particulier, que tout jeu à information parfaite possède un point d'équilibre et, par conséquent, admet une solution, c'est-à-dire il existe un couple de stratégies optimales assurant à chaque joueur un gain moyen égal à la valeur du jeu. Si un jeu à information complète ne comporte que des coups personnels, et que les deux camps s'en tiennent à leurs stratégies optimales, alors ce jeu s'achève toujours par un gain égal précisément à la valeur du jeu.

A titre d'exemple d'un jeu à information parfaite nous examinerons ci-dessous un jeu bien connu qui consiste à poser des pièces de monnaie sur une table ronde. Deux joueurs disposent à tour de rôle sur une table ronde des pièces identiques. Il est interdit de poser une pièce sur une autre. Le gagnant est celui qui le dernier arrive à trouver un endroit libre sur la table pour poser sa pièce. Il est clair que le résultat de ce jeu est connu à l'avance, il existe une stratégie qui assure incontestablement le gain à celui qui joue le premier. Plus précisément il devra placer sa première pièce juste au centre de la table et répondre ensuite à chaque trait de l'adversaire par un coup symétrique par rapport au centre. Quoi que le deuxième joueur fasse, il ne pourra en aucune manière modifier le résultat final de ce jeu. C'est pourquoi ce jeu n'a de sens que pour des joueurs ignorant la stratégie optimale. Il en est de même pour les échecs et autres jeux à information parfaite; chacun de ces jeux possède un point d'équilibre et une solution qui détermine la stratégie optimale de chaque joueur; la solution d'un jeu d'échecs n'est pas trouvée pour la simple raison que le nombre de

combinaisons des coups réalisables en est trop grand pour pouvoir construire la matrice de paiement et en définir le point d'équilibre.

§ 3. STRATÉGIES PURES ET STRATÉGIES MIXTES. RÉSOLUTION D'UN JEU AVEC DES STRATÉGIES MIXTES

Parmi les jeux finis ayant une importance pratique les jeux équilibrés sont assez rares; le cas où les valeurs inférieure et supérieure sont différentes est plus typique. En analysant les matrices de tels jeux nous sommes amenés à conclure que si chaque joueur ne peut arrêter son choix que sur une et une seule stratégie, il doit, en supposant son adversaire assez raisonnable, opter pour le principe minimax. Si l'on joue la stratégie maximin on s'assure d'avance, quelle que soit la conduite de l'adversaire, un gain égal à la valeur inférieure α du jeu. La question se pose naturellement de savoir: peut-on s'assurer un gain moyen supérieur à α en utilisant non pas une seule stratégie pure mais en alternant d'une façon aléatoire plusieurs stratégies?

De telles stratégies combinées où l'on fait alterner fortuitement plusieurs stratégies pures à une fréquence déterminée sont appelées *stratégies mixtes*.

Il est clair que chaque stratégie pure est un cas particulier de la stratégie mixte où toutes les stratégies ont une fréquence nulle à l'exception d'une qui a une fréquence égale à l'unité.

Il s'avère qu'en utilisant à la fois des stratégies pures et mixtes on peut résoudre tout jeu fini, c'est-à-dire trouver un couple de stratégies (en général mixtes) assurant aux deux joueurs, à condition qu'ils s'y tiennent, un gain égal à la valeur du jeu et, ceci étant tout joueur qui s'écarte de sa stratégie optimale ne peut que réduire son gain.

C'est sur cette assertion que repose le *théorème fondamental de la théorie des jeux*, démontré par von Neumann en 1928. La démonstration de ce théorème étant assez compliquée, nous nous bornerons seulement à son énoncé.

Tout jeu fini admet, au moins, une solution (*peut-être dans le domaine des stratégies mixtes*).

Le gain résultant d'une solution est appelé valeur du jeu. Du théorème fondamental il s'ensuit que chaque jeu fini a une valeur. Il est clair que la valeur ν d'un jeu est toujours comprise entre la valeur inférieure α et la valeur supérieure β du jeu :

$$\alpha \leqslant \nu \leqslant \beta. \tag{3.1}$$

En effet, α est le gain maximal garanti que l'on peut s'assurer en adoptant uniquement ses stratégies pures. Etant donné que toutes les stratégies pures sont un cas particulier des stratégies mixtes, en faisant appel aux premières comme aux secondes on n'empire pas, dans tous les cas, ses perspectives ; par conséquent,

$$\nu \geqslant \alpha.$$

D'une façon analogue on a relativement aux chances de l'adversaire

$$\nu \leqslant \beta,$$

ce qui entraîne (3.1).

Introduisons maintenant les notations suivantes pour les stratégies mixtes. Si, par exemple, notre stratégie mixte est composée des stratégies A_1, A_2, A_3 employées avec les fréquences p_1, p_2, p_3, où $p_1 + p_2 + p_3 = 1$, nous désignerons cette stratégie par

$$S_A = \begin{pmatrix} A_1 & A_2 & A_3 \\ p_1 & p_2 & p_3 \end{pmatrix}.$$

D'une façon analogue, la stratégie mixte de l'adversaire prend alors la forme

$$S_B = \begin{pmatrix} B_1 & B_2 & B_3 \\ q_1 & q_2 & q_3 \end{pmatrix},$$

où B_1, B_2, B_3 sont des stratégies employées avec les fréquences q_1, q_2, q_3; $q_1 + q_2 + q_3 = 1$.

Supposons maintenant qu'on ait trouvé deux stratégies mixtes optimales S_A^*, S_B^* qui sont solution du jeu. Dans le cas général, la stratégie mixte optimale du joueur donné n'englobe pas toutes les stratégies pures dont il dispose, mais seulement certaines d'entre elles. Nous appellerons stratégies « utiles » les stratégies qui font partie de la stratégie mixte optimale d'un joueur.

Il se trouve que la solution de jeu jouit encore d'une propriété remarquable suivante: *si l'un des joueurs se limite à sa stratégie mixte optimale S_A^* (S_B^*), son gain reste invariable et égal à la valeur du jeu ν indépendamment de la conduite de l'adversaire, pourvu qu'il ne s'écarte pas de ses stratégies « utiles »*. Il peut, par exemple, opter pour n'importe laquelle de ses stratégies « utiles », ou alors les alterner comme bon lui semblera.

Démontrons cette affirmation. Soit S_A^*, S_B^* la solution d'un jeu de $m \times n$. Pour fixer les idées admettons que la stratégie mixte optimale S_A^* se compose des trois stratégies « utiles » A_1, A_2, A_3; S_B^* respectivement des trois stratégies utiles B_1, B_2, B_3:

$$S_A^* = \begin{pmatrix} A_1 & A_2 & A_3 \\ p_1 & p_2 & p_3 \end{pmatrix}; \quad S_B^* = \begin{pmatrix} B_1 & B_2 & B_3 \\ q_1 & q_2 & q_3 \end{pmatrix},$$

avec $p_1 + p_2 + p_3 = 1$; $q_1 + q_2 + q_3 = 1$. On admet que si nous adoptons la stratégie S_A^*, l'adversaire peut adopter ses stratégies B_1, B_2, B_3 avec les fréquences qu'il voudra, le gain restera toujours invariable et égal à la valeur du jeu ν.

Passons à la démonstration de cette affirmation. Soient ν_1, ν_2, ν_3 les gains assurés par notre stratégie S_A^* et par les stratégies B_1, B_2 et B_3 de l'adversaire.

De la définition de la stratégie optimale il s'ensuit que si l'adversaire s'écarte de la stratégie S_B^* on aura

$$\nu_1 \geqslant \nu; \quad \nu_2 \geqslant \nu; \quad \nu_3 \geqslant \nu.$$

Voyons maintenant si l'une des valeurs ν_1, ν_2 ou ν_3 est *supérieure* à ν. Pour cela, exprimons le gain ν assuré par les stratégies optimales S_A^*, S_B^* en fonction des gains ν_1, ν_2, ν_3. Etant donné que les stratégies B_1, B_2 et B_3 figurent dans la stratégie S_B^* avec les fréquences q_1, q_2, q_3, on a

$$\nu = \nu_1 q_1 + \nu_2 q_2 + \nu_3 q_3 \quad (3.2)$$
$$(q_1 + q_2 + q_3) = 1.$$

Il est évident que si l'une au moins des valeurs ν_1, ν_2, ν_3 était plus grande que ν, leur valeur moyenne pondérée (3.2) dépasserait ν, ce qui contredit la condition. Ainsi, nous avons démontré la propriété fort importante des stratégies optimales que nous allons largement utiliser dans ce qui suit pour la résolution des jeux.

§ 4. MÉTHODES ÉLÉMENTAIRES DE RÉSOLUTION DES JEUX. JEUX DE 2×2 ET DE 2×n

Si un jeu de $m \times n$ ne présente pas de point d'équilibre, sa résolution est fort compliquée surtout pour m et n assez grands.

Parfois on arrive à simplifier ce problème, si l'on diminue préalablement le nombre de stratégies en éliminant les stratégies inutiles.

Par stratégies inutiles on entend des stratégies a) identiques et b) *a priori* désavantageuses. Examinons à titre d'exemple la matrice suivante:

A \ B	B_1	B_2	B_3	B_4
A_1	1	2	4	3
A_2	0	2	3	2
A_3	1	2	4	3
A_4	4	3	1	0

Il est facile de voir que la stratégie A_3 reproduit exactement (« double ») la stratégie A_1, l'une des deux peut donc être éliminée.

En comparant ensuite terme à terme les lignes A_1 et A_2 on constate que chaque élément de la ligne A_2 est inférieur (ou égal) à l'élément correspondant de la ligne A_1. Il est évident que nous n'adopterons jamais la stratégie A_2, celle-ci étant *a priori* désavantageuse. En rayant donc les lignes A_3 et A_2, nous pouvons réduire la matrice ci-dessus à la forme

A \ B	B_1	B_2	B_3	B_4
A_1	1	2	4	3
A_4	4	3	1	0

En poursuivant cet examen on remarque aussitôt que la stratégie B_3 est *a priori* désavantageuse pour l'adversaire. En la rayant nous réduisons la matrice à sa forme définitive

A \ B	B_1	B_2	B_4
A_1	1	2	3
A_4	4	3	0

Ainsi, en éliminant les stratégies identiques et les stratégies *a priori* désavantageuses, on peut bien ramener le jeu de 4×4 à un jeu de 2×3.

Avant de passer à la résolution d'un jeu, il importe toujours d'éliminer les stratégies identiques et les stratégies *a priori* désavantageuses.

Les jeux finis les plus simples qui se résolvent par des méthodes élémentaires sont les jeux de 2×2 et $2 \times m$.

Considérons un jeu de 2×2 dont la matrice est

A \ B	B_1	B_2
A_1	a_{11}	a_{12}
A_2	a_{21}	a_{22}

Deux cas suivants peuvent se présenter ici: 1) le jeu présente un point d'équilibre; 2) le jeu ne présente pas de point d'équilibre. Dans le premier cas la solution est immédiate: c'est le couple de stratégies dont le point d'intersection est précisément le point d'équilibre. Notons à propos que l'existence du point d'équilibre dans le jeu de 2×2 est liée à l'existence des stratégies *a priori* désavantageuses qu'il faudra éliminer *).

*) Le lecteur peut s'en convaincre en examinant quelques matrices de 2×2.

Soit un jeu sans point d'équilibre, c'est-à-dire que sa valeur inférieure n'est pas égale à la valeur supérieure: $\alpha \neq \beta$. On demande de trouver la stratégie mixte optimale du joueur A :

$$S_A^* = \begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix}.$$

Elle se caractérise par la propriété suivante: quelle que soit la stratégie adoptée de l'adversaire (à condition qu'il se limite toujours à ses stratégies « utiles »), le gain sera égal, à la valeur inférieure ν du jeu. Dans le jeu de 2×2 les deux stratégies de l'adversaire sont « utiles », sinon ce jeu aurait une solution dans le domaine des stratégies pures (le point d'équilibre). Si nous adoptons donc notre stratégie optimale $S^* = \begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix}$, l'adversaire peut répondre avec l'une de ses stratégies pures B_1 ou B_2 sans modifier d'ailleurs le gain moyen ν. On en déduit les deux équations suivantes:

$$\left.\begin{aligned} a_{11}p_1 + a_{21}p_2 = \nu, \\ a_{21}p_1 + a_{22}p_2 = \nu, \end{aligned}\right\} \qquad (4.1)$$

d'où, puisque $p_1 + p_2 = 1$, il vient

$$a_{11}p_1 + a_{21}(1 - p_1) = a_{12}p_1 + a_{22}(1 - p_1),$$

$$p_1 = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}. \qquad (4.2)$$

En portant les valeurs de p_1 et p_2 dans l'une quelconque des deux équations (4.1), on obtient la valeur ν du jeu.

La valeur du jeu étant connue, il suffit, pour déterminer la stratégie optimale $S_B^* = \begin{pmatrix} B_1 & B_2 \\ q_1 & q_2 \end{pmatrix}$ de l'adversaire, d'avoir une seule équation, par exemple l'équation

$$a_{11}q_1 + a_{12}q_2 = \nu,$$

d'où, compte tenu de l'égalité $q_1 + q_2 = 1$,

$$q_1 = \frac{\nu - a_{12}}{a_{11} - a_{12}}, \quad q_2 = 1 - q_1.$$

● EXEMPLE 1. Cherchons la solution du jeu de 2×2 de l'exemple 1 du § 1 dont la matrice est

A \ B	B_1	B_2
A_1	1	−1
A_2	−1	1

Ce jeu étant sans point d'équilibre ($\alpha = -1$; $\beta = +1$), sa solution est contenue, par conséquent, dans le domaine des stratégies mixtes:

$$S_A^* = \begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix}; \quad S_B^* = \begin{pmatrix} B_1 & B_2 \\ q_1 & q_2 \end{pmatrix}.$$

Déterminons p_1, p_2, q_1 et q_2.

Pour p_1 on a l'équation

$$1 \cdot p_1 + (-1) \cdot (1 - p_1) = (-1) \cdot p_1 + 1 \cdot (1 - p_1),$$

d'où

$$p_1 = 1/2; \; p_2 = 1/2.$$

D'une manière analogue on trouve:

$$q_1 = 1/2; \; q_2 = 1/2; \; \nu = 0.$$

Par conséquent, la stratégie optimale de chaque joueur consiste à alterner d'une façon aléatoire et avec une fréquence égale ses deux stratégies pures; le gain moyen sera alors nul.

16—0624

La conclusion que nous venons de tirer est assez évidente. Dans l'exemple qui va suivre nous examinerons un jeu plus compliqué dont la solution est loin d'être aussi apparente. C'est un exemple fort élémentaire des jeux où la « ruse » intervient. Dans les situations de conflits on fait souvent appel à des procédés qui ont pour but de tromper l'adversaire (désinformation, disposition de faux objectifs, etc.). L'exemple ci-dessous, tout élémentaire qu'il soit, est fort instructif.

● EXEMPLE 2. Soient deux cartes: un as et un deux. Le joueur A en tire une au hasard sans la faire voir à B. S'il a tiré l'as, il dit: « j'ai l'as » et demande à B de lui verser un rouble. S'il a tiré le deux, il peut alors soit A_1) dire qu'il a « tiré l'as » et demander qu'on lui verse un franc, soit A_2) avouer qu'il a le deux et verser un franc à son adversaire.

Quant à l'adversaire, si on lui paye un franc il ne peut naturellement que l'accepter. Mais si on lui demande de payer un franc, il peut soit B_1) croire A sur parole et lui payer un franc, soit B_2) demander à A de montrer sa carte. Si A a effectivement tiré l'as B lui paye deux francs, sinon c'est A qui paye deux francs à B.

On demande d'analyser ce jeu et de déterminer la stratégie optimale de chaque joueur.

● SOLUTION. Ce jeu est assez compliqué: il comprend un coup aléatoire obligatoire: le choix par A de l'une des deux cartes, et deux coups personnels qui ne sont d'ailleurs pas toujours réalisables. En effet, si A a tiré l'as, il n'a plus de coup personnel à jouer: il ne lui reste qu'à demander qu'on lui paye un franc. Dans ce cas, le coup personnel revient à B: croire ou ne pas croire (c'est-à-dire payer ou ne pas payer un franc). Si A a tiré le deux, il doit alors jouer son coup personnel: ou payer un franc à B, ou bien essayer de tromper son adversaire et lui demander un franc (bref:

« ne pas ruser » ou « ruser »). Si A choisit la première conduite, B n'a évidemment rien d'autre à faire que d'accepter un franc ; si A opte pour la deuxième conduite B est devant le dilemme suivant : croire ou ne pas croire (c'est-à-dire payer à A un franc ou demander que celui-ci lui fasse voir sa carte).

Les stratégies de chaque joueur prennent la forme de règles indiquant le comportement de chaque joueur qui a le trait.

Il est alors évident que A ne dispose que des deux stratégies suivantes :

A_1 — ruser, A_2 — ne pas ruser.

Les deux stratégies de B sont :

B_1 — croire, B_2 — ne pas croire.

Etablissons la matrice de ce jeu. Calculons pour cela le gain moyen pour chaque combinaison de stratégies.

1. A_1B_1 (A ruse, B le croit).

Si A a tiré l'as (avec la probabilité 1/2), il n'a pas joué de coup personnel ; il demande qu'on lui paye un franc et B le croit ; le gain de A en francs est 1.

Si A a tiré le deux (avec la même probabilité 1/2), il ruse, conformément à sa stratégie, et demande qu'on lui paye un franc ; B le croit et paye cette somme ; le gain de A est toujours 1.

Le gain moyen est alors :

$$a_{11} = 1/2 \cdot 1 + 1/2 \cdot 1 = 1.$$

2. A_1B_2 (A ruse, B ne le croit pas).

Si A a tiré l'as, il n'a pas de coup personnel à jouer ; il demande qu'on lui paye un franc ; B, selon sa stratégie, ne le croit pas et, après avoir vu la carte de son adversaire, lui paye deux francs (le gain de A est alors $+2$).

Si A a tiré le deux, il demande toujours (selon sa stratégie) que l'adversaire lui paye un franc ; B (selon sa stratégie aussi) ne le croit pas ; A paye donc à B deux francs

(le gain de A est -2). Le gain moyen est cette fois-ci :

$$a_{12} = 1/2 \cdot (+2) + 1/2 \cdot (-2) = 0.$$

3. A_2B_1 (A ne ruse pas, B le croit).

Si A a tiré l'as, il demande qu'on lui paye un franc ; B, conformément à sa stratégie, lui paye cette somme ; le gain de A est égal à $+1$. Si A a tiré le deux, il paye alors un franc à B qui n'a rien d'autre à faire que d'accepter cette somme (le gain de A est alors égal à -1). Le gain moyen est :

$$a_{21} = 1/2 \cdot (+1) + 1/2 \cdot (-1) = 0.$$

4. A_2B_2 (A ne ruse pas, B ne le croit pas).

Si A a tiré l'as, il demande qu'on lui paye un franc ; B ne le croit pas et, après avoir vu sa carte, paye à son adversaire deux francs (le gain de A est alors $+2$).

Si A a tiré le deux, il paye tout de suite à B un franc, B l'accepte (le gain de A est -1).

Le gain moyen représente alors :

$$a_{22} = 1/2 \cdot (+2) + 1/2 \cdot (-1) = 1/2.$$

Formons maintenant la matrice de ce jeu :

A \ B	B_1 (croire)	B_2 (ne pas croire)
A_1 (ruser)	1	0
A_2 (ne pas ruser)	0	1/2

Celle-ci n'a pas de point d'équilibre. La valeur inférieure du jeu est $\alpha = 0$, la valeur supérieure $\beta = 1/2$. Cherchons la solution du jeu dans le domaine des stratégies mixtes. La formule (4.2) donne :

$$p_1 = \frac{1/2}{1+1/2} = 1/3 ; \quad p_2 = 2/3 ; \quad S_A^* = \begin{pmatrix} A_1 & A_2 \\ 1/3 & 2/3 \end{pmatrix},$$

c'est-à-dire A doit, dans un tiers des cas, adopter sa première stratégie (ruser) et dans les deux tiers des cas sa seconde stratégie (ne pas ruser). Il s'assurera alors un gain égal à la valeur moyenne du jeu, soit

$$\nu = 1/3.$$

La valeur $\nu = 1/3$ montre qu'à ces conditions ce jeu est avantageux pour A et désavantageux pour B. En choisissant sa stratégie optimale, A peut se garantir à coup sûr un gain moyen positif.

Remarquons que si A adoptait sa stratégie la plus prudente (stratégie maximin ; dans le cas considéré les deux stratégies A_1 et A_2 sont maximin), son gain moyen serait nul. Ainsi, en choisissant sa stratégie mixte, A prend l'avantage sur B conformément aux règles du jeu.

Déterminons maintenant la stratégie optimale de B. On a :

$$q_1 \cdot 1 + q_2 \cdot 0 = 1/3 ; \; q_1 = 1/3 ; \; q_2 = 2/3,$$

d'où $S_B^* = \begin{pmatrix} B_1 & B_2 \\ 1/3 & 2/3 \end{pmatrix}$, c'est-à-dire le joueur B doit dans un tiers des cas, croire A et lui payer un franc sans lui demander de montrer ses cartes, et, dans les deux tiers des cas, vérifier les cartes de l'adversaire. Ses pertes ne dépasseront pas alors 1/3 en moyenne pour chaque coup joué. Mais s'il utilisait sa stratégie pure minimax B_2 (ne pas ruser), il perdrait alors 1/2 en moyenne pour chaque coup joué.

La solution d'un jeu de 2×2 admet une interprétation géométrique fort simple. Soit un jeu de 2×2 dont la matrice est

A \ B	B_1	B_2
A_1	a_{11}	a_{12}
A_2	a_{21}	a_{22}

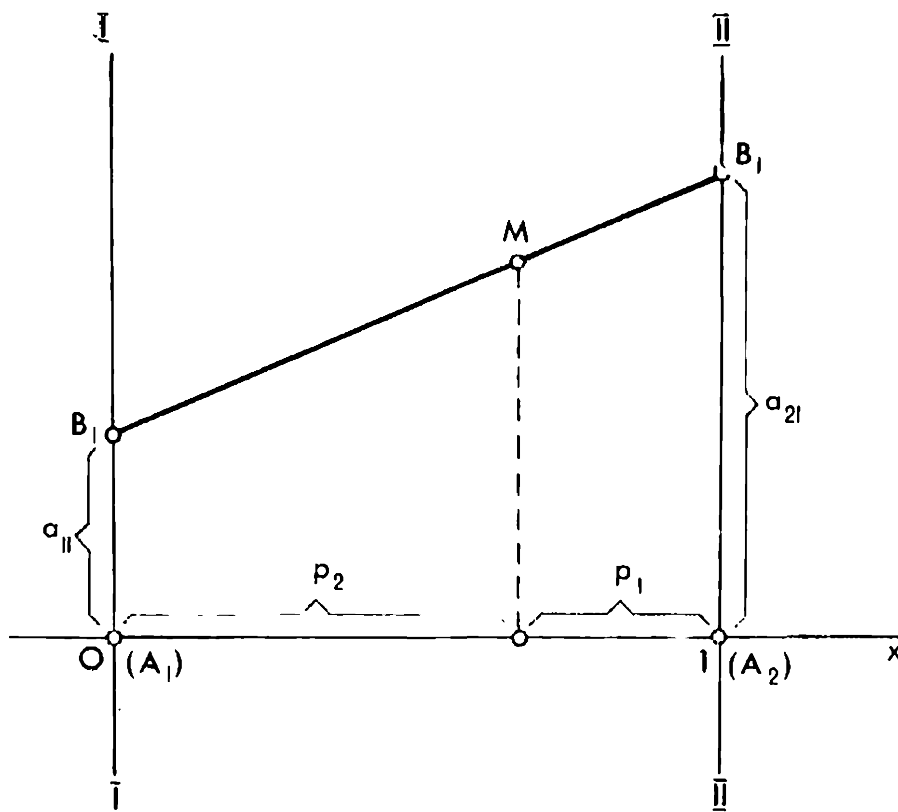


Fig. 1

Choisissons un intervalle de l'axe des abscisses de longueur 1 (fig. 1). L'extrémité gauche (le point d'abscisse $x = 0$) représentera la stratégie A_1; l'extrémité droite ($x = 1$) la stratégie A_2. Elevons à partir des points A_1 et A_2 deux perpendiculaires à l'axe des abscisses que nous désignerons respectivement par axe $I - I$ et axe $II - II$. Portons sur l'axe $I - I$ les gains correspondant à la stratégie A_1 et sur l'axe $II - II$ les gains correspondant à la stratégie A_2. Examinons maintenant la stratégie B_1 de l'adversaire; elle nous donne deux points sur les axes $I - I$ et $II - II$ d'ordonnées a_{11} et a_{21} respectivement. Joignons ces points par une droite. Il est alors évident que si à la stratégie B_1 de l'adversaire nous répondons par la stratégie mixte $S_A = \begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix}$, notre gain moyen, égal dans ce cas à $a_{11}p_1 + a_{12}p_2$, sera représenté par le point M de la droite B_1B_1;

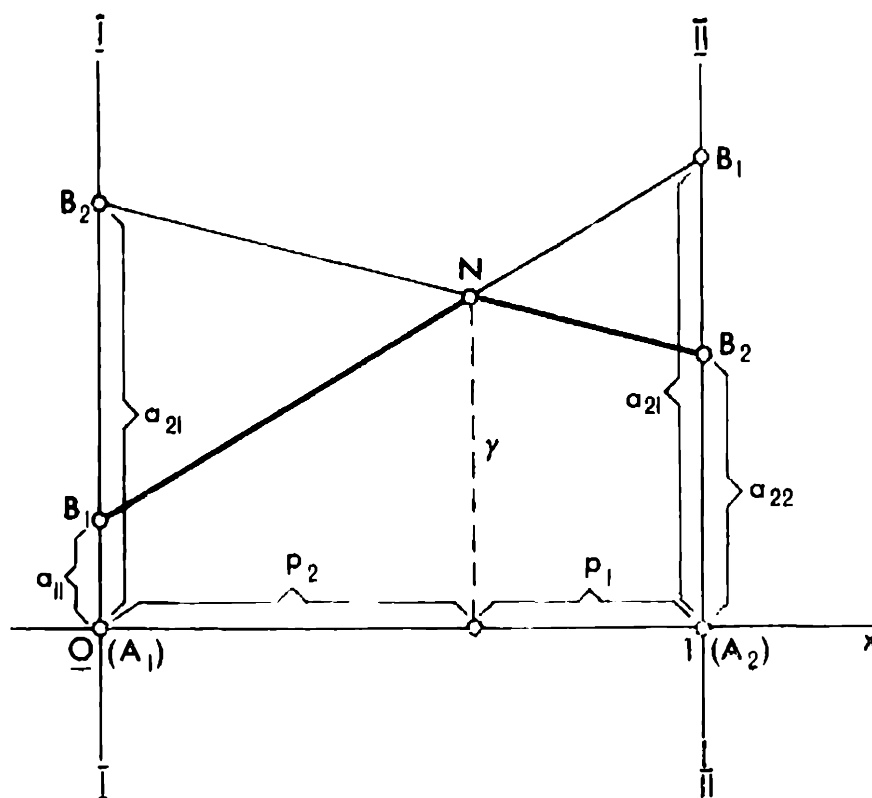


Fig. 2

d'abscisse égale à p_2. Nous conviendrons d'appeler « stratégie B » la droite B_1B_1 représentant le gain correspondant à la stratégie B_1.

Il est aisé de voir que la stratégie B_2 (fig. 2) peut être construite d'une manière analogue.

Il nous faut déterminer la stratégie optimale S_A^*, c'est-à-dire une stratégie telle que le gain minimal (quel que soit le comportement de B) se transforme en gain maximal. Traçons pour cela la *limite inférieure du gain* pour les stratégies B_1, B_2, c'est-à-dire la ligne polygonale B_1NB_2 en traits gras sur la fig. 2. Cette limite inférieure représente le gain minimal de A quelles que soient les stratégies mixtes de ce joueur; le point N, où le gain minimal atteint son maximum est donc la solution et la valeur du jeu considéré. On s'assure aisément que l'ordonnée du point N représente la valeur

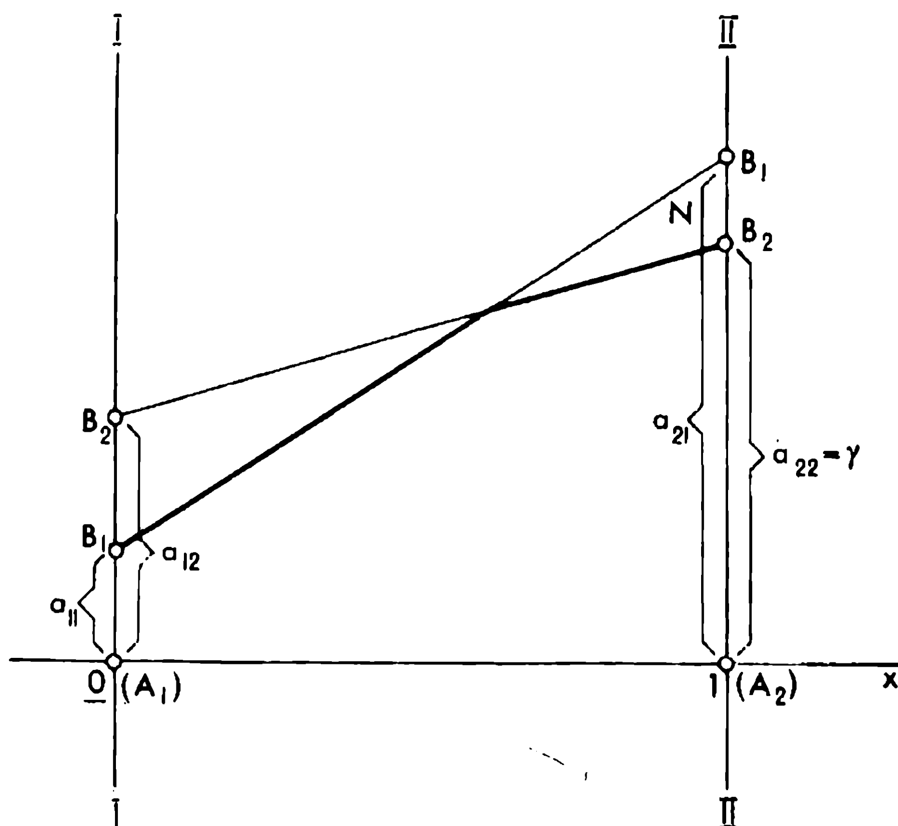


Fig. 3

du jeu ν et son abscisse est égale à p_2, fréquence de la stratégie A_2 dans la stratégie optimale mixte S_A^*.

Dans le cas considéré la solution du jeu est donnée par le point d'intersection des stratégies. Néanmoins, ce n'est pas une règle générale ; ainsi, on a représenté sur la fig. 3 un cas où la solution n'est pas donnée par le point d'intersection des stratégies, mais par les stratégies pures A_2 et B_2 des deux joueurs et où la valeur du jeu est $\nu = a_{22}$.

Dans le cas considéré la matrice a un point d'équilibre et la stratégie A_1 est *a priori* désavantageuse, car elle assure un gain inférieur à A_2 quelle que soit la stratégie pure de l'adversaire.

Dans le cas où l'adversaire est contraint de jouer une stratégie *a priori* désavantageuse, l'interprétation géométrique prend la forme représentée sur la fig. 4,

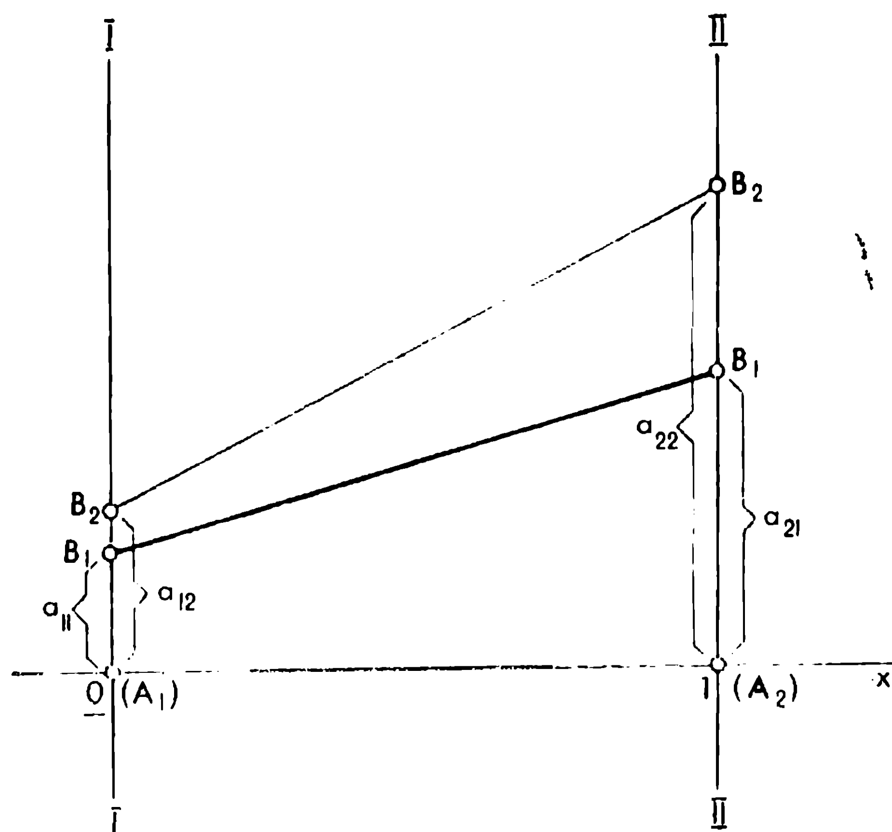


Fig. 4

Dans ce cas la limite inférieure du gain se confond avec la stratégie B_1; la stratégie B_2 est a *priori* désavantageuse pour l'adversaire.

L'interprétation géométrique permet de rendre plus concrètes les valeurs inférieure et supérieure du jeu (fig. 5). Interprétons géométriquement les jeux de 2×2 des exemples 1 et 2 (fig. 6 et 7).

Tout jeu de 2×2 peut être résolu, comme nous l'avons montré plus haut, par des procédés élémentaires. Tout jeu de $2 \times n$, où nous n'avons que deux stratégies et l'adversaire un nombre arbitraire de stratégies, peut être résolu d'une manière absolument analogue.

Soient A_1, A_2 nos deux stratégies et $B_1, B_2, \ldots, B_n$, les n stratégies de l'adversaire. La matrice a_{ij} est donnée; elle comprend deux lignes et n colonnes. Comme dans le

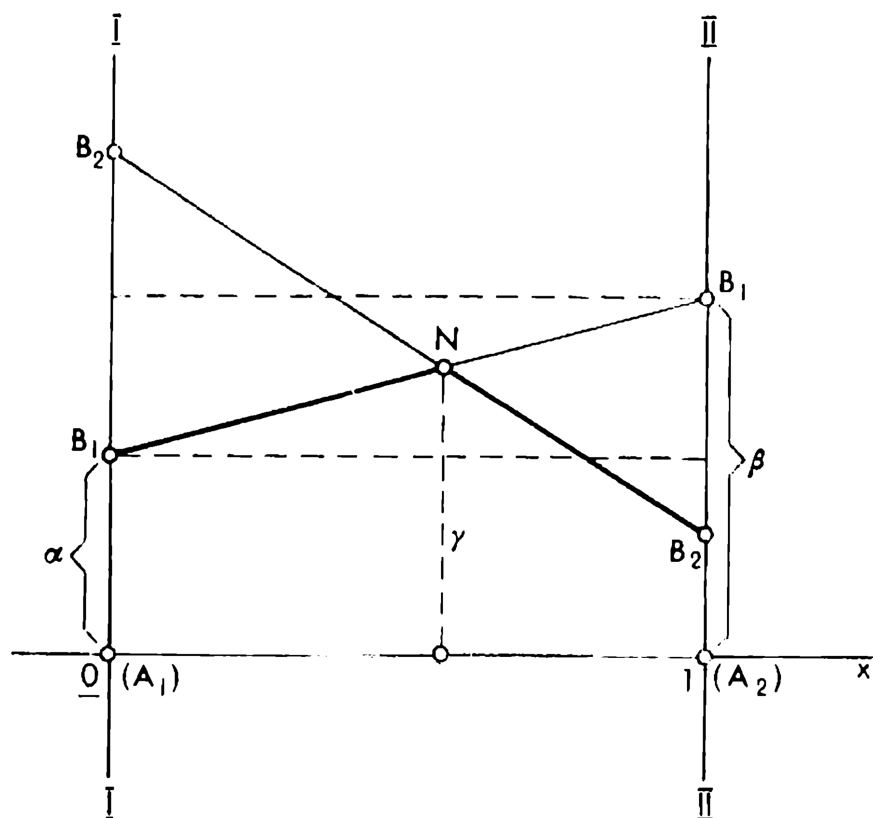


Fig. 5

cas précédent de deux stratégies, donnons à ce problème une interprétation géométrique; les n stratégies de l'adversaire seront représentées (fig. 8) par n droites. Traçons la limite inférieure du gain (la ligne polygonale B_1MNB_2) et déterminons sur elle le point N d'ordonnée maximale. Ce point représente donc la solution du jeu considéré (la stratégie $S_A^* = \begin{pmatrix} A_1 & A_2 \\ p_1 & p_2 \end{pmatrix}$); l'ordonnée du point N est égale à la valeur du jeu ν et son abscisse à la fréquence p_2 de la stratégie A_2.

Dans le cas considéré la stratégie optimale de l'adversaire est une combinaison de deux stratégies « utiles » B_2 et B_4 qui se coupent au point N. La stratégie B_3 est *a priori* désavantageuse et B_1 l'est aussi si nous jouons notre stratégie optimale S_A^*. Si A adopte sa stratégie optimale il s'assu-

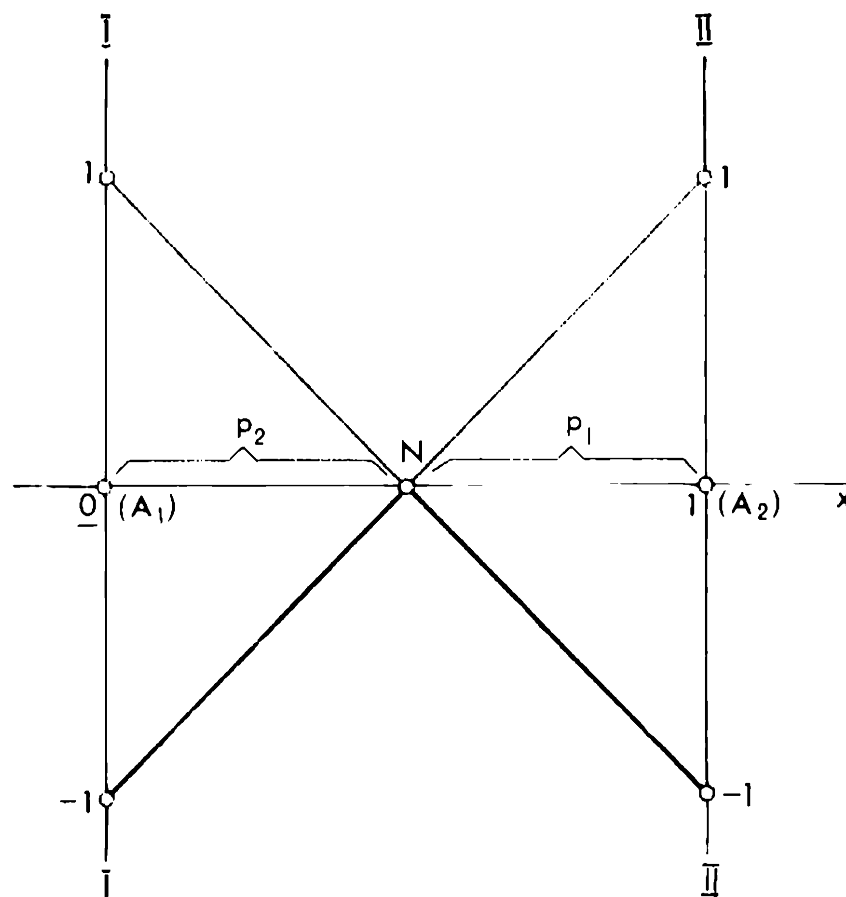


Fig. 6

rera un gain constant quelle que soit la stratégie « utile » de B. Mais si ce dernier joue ses stratégies B_1 ou B_2 le gain de A changera de valeur.

Dans la théorie des jeux on démontre que tout jeu fini de $m \times n$ admet une solution où le nombre de stratégies « utiles » de l'un ou de l'autre joueur ne dépasse pas le plus petit des nombres m et n. D'où il s'ensuit en particulier que tout jeu de $2 \times m$ admet toujours une solution composée au plus de deux stratégies « utiles » de chaque participant.

En faisant appel à une interprétation géométrique on peut donner un procédé fort élémentaire de résolution de tout jeu de $2 \times m$. Déterminons sur le dessin même un couple de stratégies « utiles » B_j et B_k se coupant au point N (si au point N se coupent plusieurs stratégies, nous en choi-

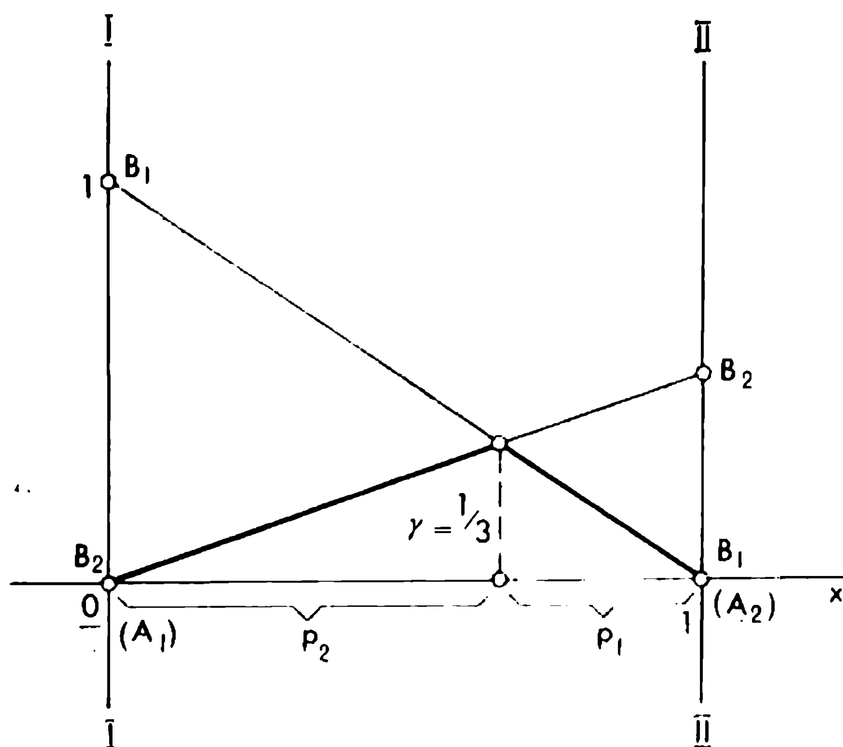


Fig. 7

sissons deux arbitrairement). Comme nous l'avons vu plus haut, si A joue sa stratégie optimale, son gain ne dépend pas de la combinaison dans laquelle B emploie ses stratégies « utiles ». Par conséquent,

$$\left.\begin{aligned} p_1 a_{1j} + p_2 a_{2j} = \nu, \\ p_1 a_{1k} + p_2 a_{2k} = \nu. \end{aligned}\right\}$$

De ces deux équations et de la condition $p_2 = 1 - p_1$, on détermine p_1, p_2 et la valeur du jeu ν.

Une fois la valeur du jeu connue, on trouve aussitôt la stratégie optimale $S_B^* = \begin{pmatrix} B_j & B_k \\ q_j & q_k \end{pmatrix}$ de B.

Pour cela il faut résoudre, par exemple, l'équation

$$q_j a_{1j} + q_k a_{1k} = \nu,$$

où

$$q_j + q_k = 1.$$

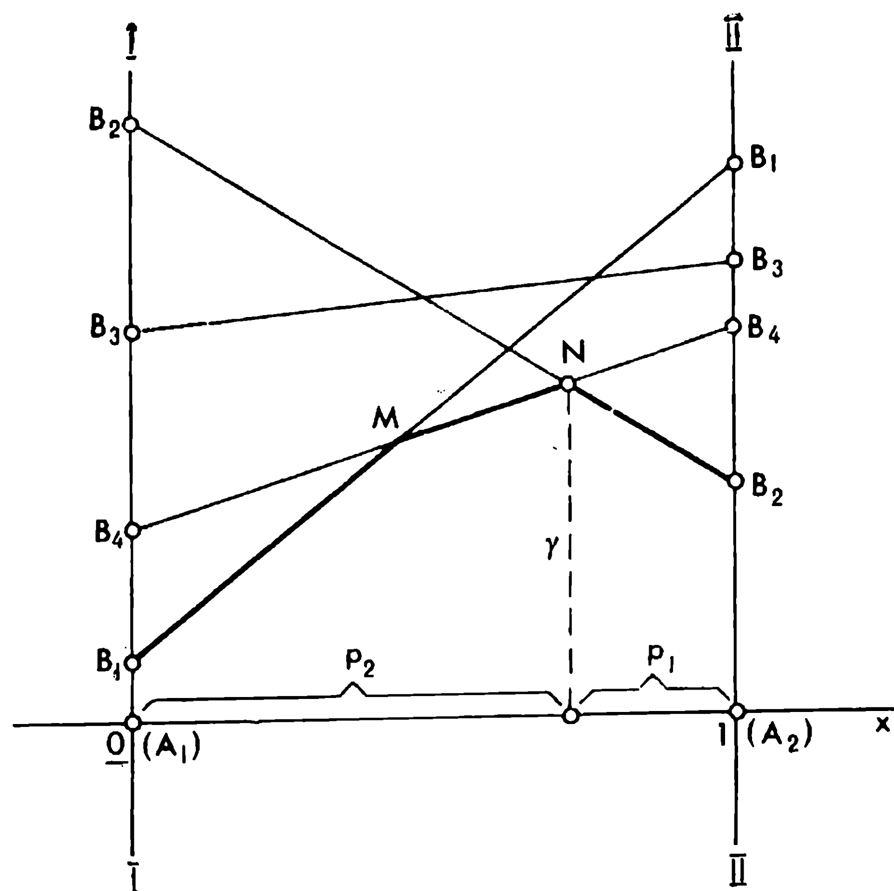


Fig. 8

Il est aisé de voir que lorsque nous disposons de m stratégies et l'adversaire de deux stratégies seulement, le problème se résout d'une façon absolument analogue ; il suffit de remarquer qu'en changeant le signe du gain on transforme A de « gagnant » en « perdant ». On peut d'ailleurs résoudre ce jeu sans recourir au changement de signe ; on commence à résoudre ce problème immédiatement pour B, en construisant non pas la limite inférieure mais la limite supérieure du gain (fig. 9). On trouve ensuite sur celle-ci le point N d'ordonnée minimale qui représente exactement la valeur du jeu v.

Examinons et résolvons quelques jeux de 2×2 et $2 \times m$ qui sont des modèles simplifiés de jeux ayant une importance pratique.

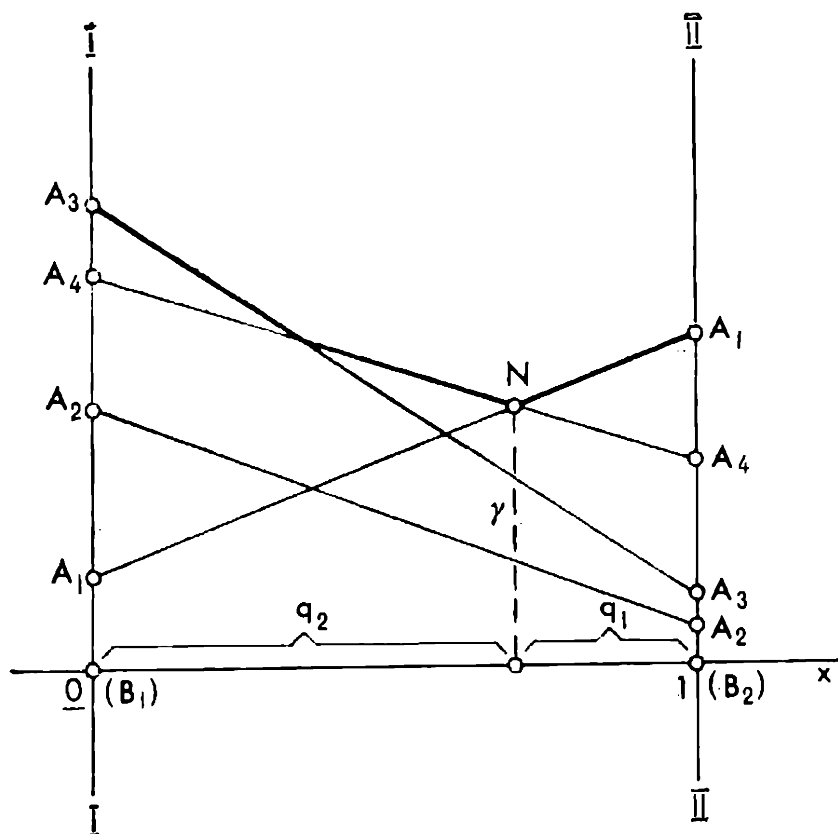


Fig. 9

● EXEMPLE 3. *A* attaque avec deux bombardiers *I* et *II* la zone de protection de l'adversaire *B*; le bombardier *I* précède le *II*. L'un d'eux — on ne sait pas d'avance lequel — porte la bombe, l'autre est chargé de l'escorte. Dans la zone ennemie, les bombardiers sont interceptés par des chasseurs de *B*. Les bombardiers sont équipés de canons de cadence de tir différente. Si le chasseur ennemi attaque l'avion d'escorte *II* il ne s'expose qu'aux canons de ce dernier; s'il attaque le premier avion il s'expose aux tirs des deux bombardiers. La probabilité d'atteindre le chasseur est 0,3 dans le premier cas, et 0,7 dans le second.

Si le chasseur n'est pas abattu par les bombardiers, il atteint la cible qu'il s'est choisie avec la probabilité 0,6. Le but des bombardiers est de lancer leur bombe sur un objectif; celui du chasseur est de s'y opposer, c'est-à-dire

d'abattre le bombardier portant la bombe. On demande de définir les stratégies optimales des deux parties:

a) pour A : quel bombardier doit porter la bombe?

b) pour B : quel bombardier faut-il attaquer?

● SOLUTION. C'est un cas assez simple de jeu de 2×2; le gain est la probabilité de ne pas atteindre le bombardier chargé d'une bombe.

Nos stratégies:

A_1 — le bombardier *I* porte la bombe;

A_2 — le bombardier *II* porte la bombe.

Les stratégies de l'adversaire:

B_1 — attaquer le bombardier *I*;

B_2 — attaquer le bombardier *II*.

Formons la matrice du jeu considéré, c'est-à-dire déterminons le gain moyen pour chaque combinaison de stratégies.

1. A_1B_1 (*I* porte la bombe et c'est lui qui est attaqué).

Le porteur ne sera pas atteint si les bombardiers abattent le chasseur ou bien ils n'abattent pas le chasseur mais ce dernier ne pourra pas atteindre son but.

$$a_{11} = 0,7 + 0,3 \cdot 0,4 = 0,82.$$

2. A_2B_1 (*II* porte la bombe, on attaque *I*)

$$a_{21} = 1.$$

3. A_1B_2 (*I* porte la bombe, on attaque *II*)

$$a_{12} = 1.$$

4. A_2B_2 (*II* porte la bombe et on attaque *II*)

$$a_{22} = 0,3 + 0,7 \cdot 0,4 = 0,58.$$

La matrice du jeu considéré prend alors la forme :

A \ B	B_1	B_2
A_1	0,82	1
A_2	1	0,58

La valeur inférieure est 0,82, la valeur supérieure 1. La matrice n'a pas de point d'équilibre ; cherchons la solution dans le domaine des stratégies mixtes.

On a :

$$p_1 \cdot 0{,}82 + p_2 \cdot 1 = v ;$$

$$p_1 \cdot 1 + p_2 \cdot 0{,}58 = v ;$$

$$p_2 = 1 - p_1.$$

D'où

$$p_1 = 0{,}7 ; \; p_2 = 0{,}3.$$

Notre stratégie optimale est

$$S_A^* = \begin{pmatrix} A_1 & A_2 \\ 0{,}7 & 0{,}3 \end{pmatrix},$$

c'est-à-dire nous devons choisir pour porteur le plus souvent possible le bombardier *I*. La valeur du jeu est alors

$$v = 0{,}874.$$

Déterminons maintenant les fréquences q_1 et q_2 des stratégies B_1 et B_2 dans la stratégie optimale S_B^* de l'adversaire. On a :

$$q_1 \cdot 0{,}82 + q_2 \cdot 1 = 0{,}874 ;$$

d'où

$$q_2 = 1 - q_1,$$

$$q_1 = 0{,}7 ; \quad q_2 = 0{,}3,$$

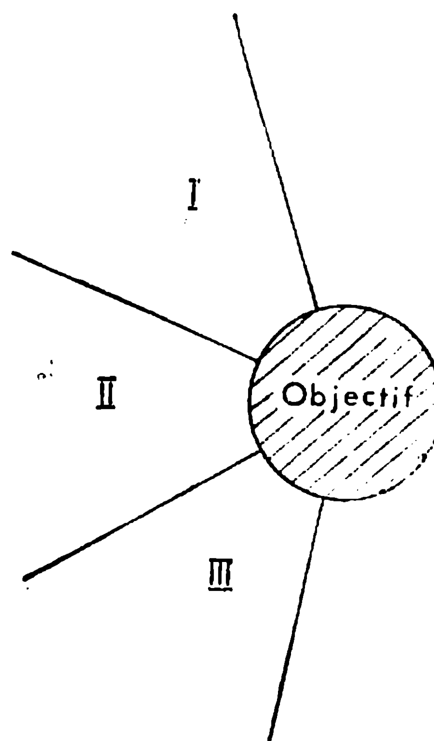


Fig. 10

c'est-à-dire la stratégie optimale de l'adversaire est

$$S_B^* = \begin{pmatrix} B_1 & B_2 \\ 0,7 & 0,3 \end{pmatrix}.$$

● EXEMPLE 4. A attaque un objectif, la partie B le défend. A dispose pour cela de deux avions, B de trois pièces de D.C.A. Chaque avion possède un moyen puissant de destruction, donc, pour détruire l'objectif il suffit qu'un seul avion franchisse la défense antiaérienne de l'adversaire. Les avions de A peuvent attaquer l'objectif à partir de trois directions différentes : I, II, III (fig. 10).

B peut, à son tour, disposer ses canons de façon à protéger n'importe quelle direction menacée ; chaque canon ne tient sous le feu qu'un secteur d'approche correspondant exclusivement à la direction choisie. Chaque canon ne peut tirer que sur un seul avion avec une probabilité de l'atteindre égale à 1. A ne sait pas où sont installés les canons ;

B ne sait pas d'où arriveront les avions. Le but de A est d'atteindre l'objectif, celui de B de l'en empêcher. Trouver la solution de ce jeu.

● SOLUTION. Ce jeu est un jeu de 2×3. Le gain est la probabilité d'atteindre l'objectif. Nos stratégies possibles sont :

A_1 — attaquer l'objectif avec deux avions dans deux directions différentes,

A_2 — attaquer celui-ci avec deux avions à la fois dans une seule direction.

Les stratégies possibles de l'adversaire sont :

B_1 — placer un canon dans chaque direction menacée ;

B_2 — placer deux canons dans une direction quelconque, un seul sur une autre, et ne pas couvrir la troisième ;

B_3 — placer les trois canons dans une direction quelconque et ne pas couvrir les autres.

Formons la matrice du jeu considéré.

1. A_1B_1 (les avions attaquent l'objectif dans les deux directions différentes ; les canons couvrent toutes les trois directions).

Il est aisé de comprendre que dans ce cas aucun avion n'atteindra l'objectif attaqué ;

$$a_{11} = 0.$$

2. A_2B_1 (les deux avions attaquent dans une direction quelconque ; les canons tiennent sous le feu les trois directions menacées).

On comprend aisément qu'un avion atteindra l'objectif fixé sans être inquiété :

$$a_{21} = 1.$$

3. A_1B_2 (les avions attaquent l'objectif séparément ; l'adversaire protège deux directions seulement).

La probabilité qu'au moins un avion force la défense ennemie est égale à la probabilité que l'un d'eux puisse trouver la direction non protégée:

$$a_{12} = 2/3.$$

4. A_2B_2 (les deux avions attaquent l'objectif ensemble suivant la même direction; l'adversaire protège une direction avec deux canons et une autre avec un seul canon, c'est-à-dire en fait il ne couvre qu'une seule direction).

La probabilité qu'au moins un avion franchisse la défense ennemie est égale à la probabilité que les deux avions trouvent la direction non protégée:

$$a_{22} = 2/3.$$

5. A_1B_3 (les deux avions attaquent l'objectif dans des directions différentes; l'adversaire protège une seule direction avec ses trois canons).

$$a_{13} = 1.$$

6. A_2B_3 (les deux avions attaquent l'objectif ensemble; l'adversaire place tous ses trois canons dans une direction). Pour atteindre l'objectif, les avions doivent voler dans une direction non protégée:

$$a_{23} = 2/3.$$

La matrice du jeu considéré est

A \ B	B_1	B_2	B_3
A_1	0	2/3	1
A_2	1	2/3	2/3

On constate aisément que la stratégie B_3 est *a priori* désavantageuse en comparaison de B_2 (on aurait pu s'en rendre compte dès le début). En éliminant donc la stratégie B_3 on obtient un jeu de 2×2:

A \ B	B_1	B_2
A_1	0	2/3
A_2	1	2/3

Cette matrice est à point d'équilibre: la valeur inférieure du jeu 2/3 se confond avec la valeur supérieure.

On remarque aussitôt que la stratégie A_1 est *a priori* désavantageuse pour nous (A). Conclusion: les deux parties A et B doivent toujours utiliser leurs stratégies pures A_2 et B_2, c'est-à-dire nous devons attaquer l'objectif fixé avec deux avions à la fois dans une direction aléatoirement choisie; l'adversaire quant à lui doit disposer ses canons comme suit: deux canons dans une direction et le troisième dans une autre; le choix des directions protégées doit également être aléatoire (comme on le voit, ici les « stratégies pures » revêtent un caractère aléatoire). En adoptant ces stratégies optimales, nous nous assurerons toujours un gain moyen constant égal à 2/3 (c'est-à-dire, l'objectif sera atteint avec la probabilité 2/3).

Observons que la solution obtenue n'est pas unique; en dehors des stratégies pures, il existe toute une série de stratégies mixtes de A qui sont également optimales, de $p_1 = 0$ à $p_1 = 1/3$ (fig. 11). On s'assure aisément que le même gain moyen de 2/3 peut être obtenu si l'on adopte les stratégies A_1 et A_2 dans la proportion 1/2 et 2/3.

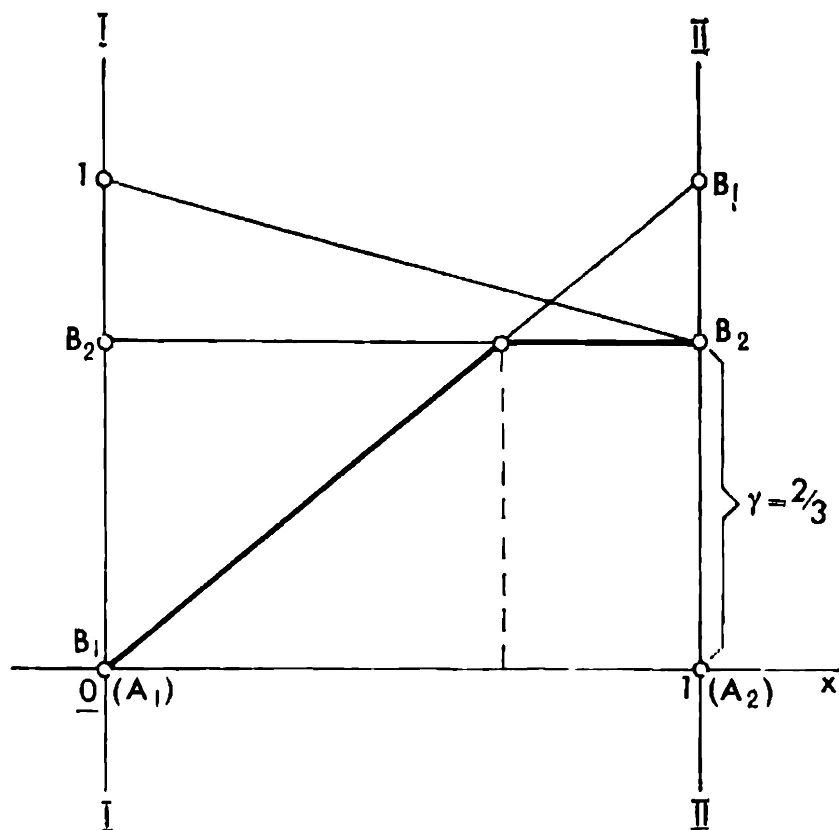


Fig. 11.

● EXEMPLE 5. Traiter l'exemple précédent avec quatre directions possibles pour attaquer l'objectif et quatre canons chez l'adversaire.

● SOLUTION. Nous avons comme précédemment deux stratégies possibles :

A_1 — attaquer avec nos deux avions dans des directions différentes,

A_2 — attaquer avec nos deux avions ensemble.

L'adversaire, lui, a cinq stratégies possibles :

B_1 $(1 + 1 + 1 + 1)$ — placer un canon dans chaque direction menacée ;

B_2 $(2 + 2)$ — placer deux canons dans deux directions différentes ;

B_3 $(2 + 1 + 1)$ — placer deux canons dans une direction quelconque, deux dans deux autres et laisser une direction sans défense ;

B_4 (3 + 1) — placer trois canons dans une direction, un dans une autre et laisser deux directions sans défense;

B_5 (4) — placer les quatre canons dans une seule direction.

Les stratégies B_4, B_5 sont à rejeter comme étant *a priori* désavantageuses. Comme dans l'exemple précédent, formons la matrice du jeu considéré:

A \ B	B_1 (1+1+1+1)	B_2 (2+2)	B_3 (2+1+1)
A_1	0	5/6	1/2
A_2	1	1/2	3/4

La valeur inférieure du jeu est 1/2, la valeur supérieure 3/4.

La matrice en question est sans point d'équilibre; la solution est donc comprise dans le domaine des stratégies mixtes. Faisons appel à une interprétation géométrique de ce jeu (fig. 12), pour déterminer les stratégies « utiles » B_1 et B_2 de l'adversaire.

Les fréquences p_1 et p_2 s'obtiennent des équations:

$$p_1 \cdot 0 + (1 - p_1) \cdot 1 = \nu;$$
$$p_1 \cdot 5/6 + (1 - p_1) \cdot 1/2 = \nu,$$

d'où

$$p_1 = 3/8;$$
$$p_2 = 5/8;$$
$$\nu = 5/8,$$

c'est-à-dire notre stratégie optimale est

$$S_A^* = \begin{pmatrix} A_1 & A_2 \\ 3/8 & 5/8 \end{pmatrix}.$$

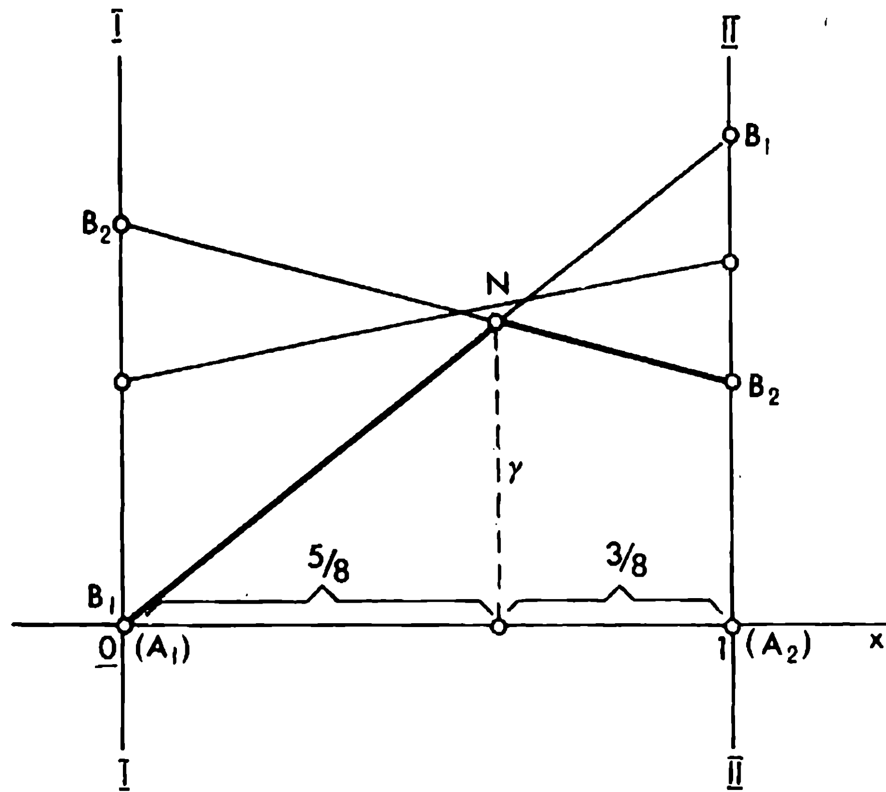

I
II
B1
B2
N
B2
γ
5/8
3/8
B1
0 (A1)
1 (A2)
x
I
II

Fig. 12

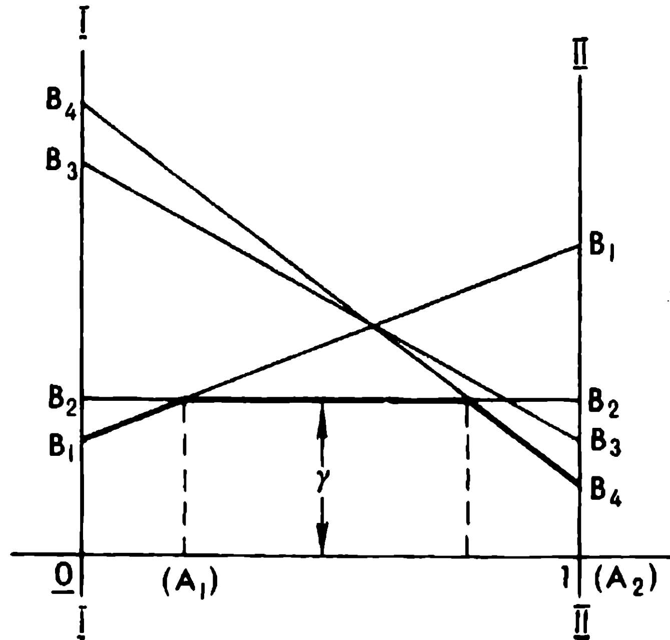

I
II
B4
B3
B1
B2
B2
B1
B3
B4
γ
0
(A1)
1 (A2)
I
II

Fig. 13

En adoptant cette stratégie nous nous assurons un gain moyen égal à 5/8. Après avoir déterminé la valeur du jeu $v = 5/8$, il est aisé de trouver les fréquences q_1 et q_2 des stratégies « utiles » de l'adversaire:

$$q_1 \cdot 0 + (1 - q_1) \cdot 5/6 = 5/8,$$
$$q_1 = 1/4; \; q_2 = 3/4.$$

La stratégie optimale de l'adversaire est donc:

$$S_B^* = \begin{pmatrix} B_1 & B_2 \\ 1/4 & 3/4 \end{pmatrix}.$$

● EXEMPLE 6. A dispose de deux stratégies A_1 et A_2, B, de quatre stratégies suivantes B_1, B_2, B_3 et B_4. La matrice de ce jeu a la forme

B / A	B_1	B_2	B_3	B_4
A_1	3	4	10	12
A_2	8	4	3	2

On demande de trouver la solution de ce jeu.

● SOLUTION. Les valeurs inférieure et supérieure du jeu considéré sont respectivement 0,3 et 0,4. En faisant appel à une interprétation géométrique (fig. 13) nous constatons aisément que les stratégies utiles du joueur B sont B_1 et B_2 ou bien B_2 et B_4. Quant au joueur A, il dispose d'une infinité de stratégies mixtes optimales: p_1 peut varier de 1/5 à 4/5 dans une stratégie optimale. La valeur du jeu est $v = 4$. La stratégie optimale pure du joueur B est B_2.

§ 5. MÉTHODES GÉNÉRALES DE RÉSOLUTION DES JEUX FINIS

Jusqu'ici nous n'avons examiné que les jeux les plus élémentaires du type $2 \times n$ qui se résolvent facilement et admettent une interprétation géométrique simple.

Dans le cas général la résolution d'un jeu de $m \times n$ est assez compliquée. Les difficultés ne portent pas sur le principe (qui est le même quel que soit m) du problème, elles tiennent surtout aux calculs encombrants qui sont parfois pratiquement irréalisables lorsque m et n sont grands.

Illustrons cela par l'exemple d'un jeu de $3 \times n$. Commençons par son interprétation géométrique, mais cette fois-ci, dans l'espace. Représentons nos trois stratégies A_1, A_2 et A_3 par trois points dans le plan xOy: le premier se confond avec l'origine des coordonnées (fig. 14), le deuxième et le troisième se trouvent respectivement sur les axes Ox et Oy à la distance 1 de l'origine.

Par les points A_1,A_2 et A_3 menons les axes $I - I$, $II - II$ et $III - III$ perpendiculaires au plan xOy. Portons sur l'axe $I - I$ les gains correspondant à la stratégie A_1, sur $II - II$ et $III - III$ les gains assurés respectivement par les stratégies A_2 et A_3. Chaque stratégie B_j de l'adversaire prendra alors la forme d'un plan découpant sur les axes $I - I$, $II - II$ et $III - III$ des segments correspondant aux gains assurés respectivement par les stratégies A_1, A_2, A_3 et par la stratégie B_j. En construisant ainsi toutes les stratégies de l'adversaire, on obtient une famille de plans disposés au-dessus du triangle $A_1A_2A_3$ (fig. 15). On peut construire pour cette famille aussi la limite inférieure du gain, comme dans le cas du jeu de $2 \times n$ déjà considéré, et déterminer ensuite sur cette limite le point N de hauteur maximale par rapport au plan xOy. Cette hauteur représentera la valeur du jeu v considéré. Les fréquences p_1, p_2, p_3 des stratégies A_1, A_2, A_3 dans la stratégie optimale S_A^* se détermineront par les coordonnées (x, y) du point N, à

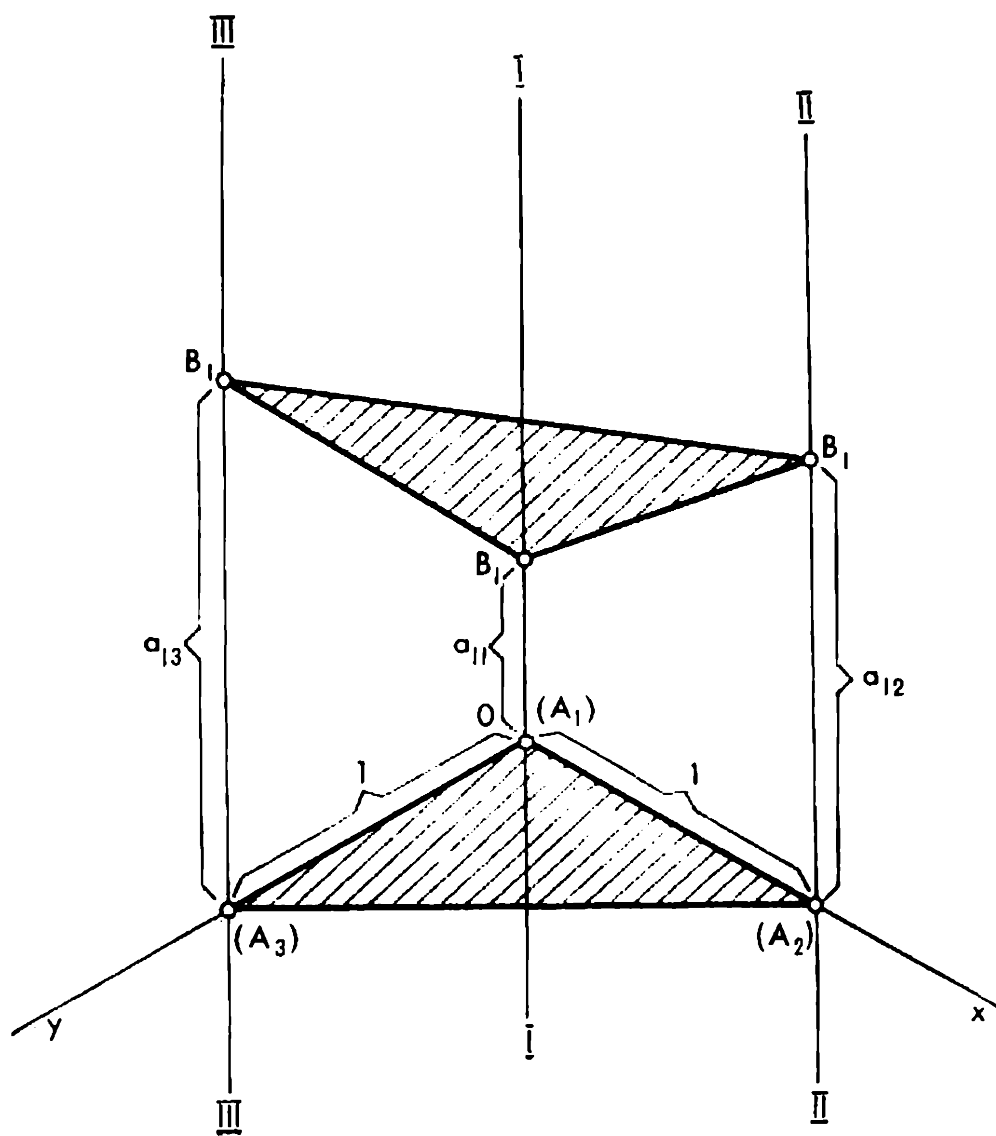


Fig. 14

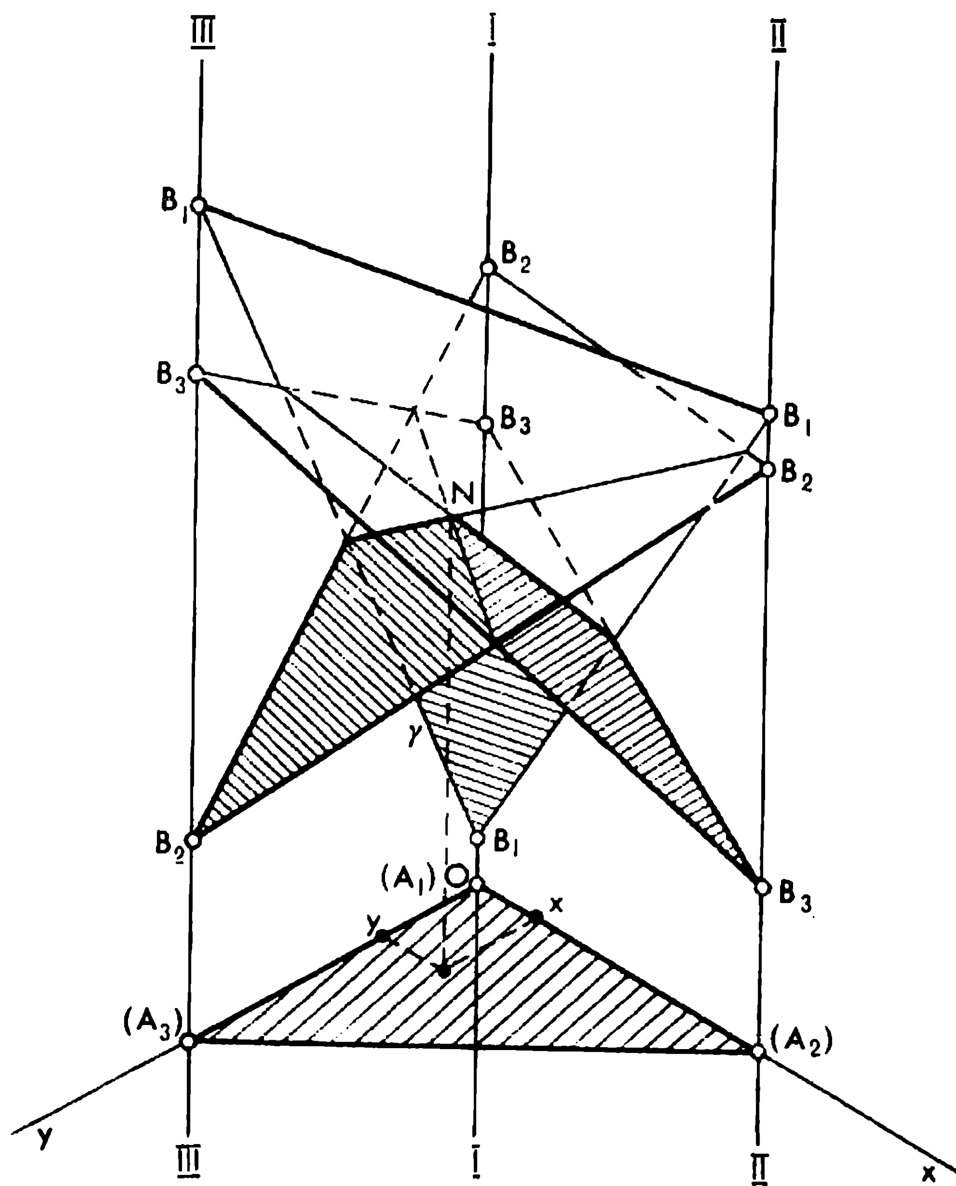


Fig. 15

savoir:

$$p_2 = x\,;\ p_3 = y\,;\ p_1 = 1 - p_2 - p_3.$$

Néanmoins, une telle représentation géométrique même dans le cas d'un jeu de $3 \times n$ est bien difficile à réaliser et nécessite beaucoup de temps et d'imagination. Dans le cas général cette construction géométrique se trouve reportée dans un espace m-dimensionnel et cesse d'être concrète, bien que l'emploi de la terminologie géométrique puisse être utile. Dans la résolution pratique des jeux de $m \times n$ les méthodes analytiques de calcul s'avèrent meilleures par rapport aux analogies géométriques, d'autant plus que ces méthodes sont les seules qui passent sur ordinateur.

Toutes ces méthodes se ramènent au fond à la résolution d'un problème par le procédé des approximations successives, mais la mise en ordre de la succession des opérations permet de former un algorithme qui fournit le plus vite possible la solution recherchée.

Nous exposerons brièvement ci-dessous une méthode de résolution numérique des jeux de $m \times n$, dite méthode de « programmation linéaire ».

Commençons par énoncer la position générale du problème. Soit un jeu de $m \times n$ où le joueur A dispose de m stratégies $A_1, A_2, \ldots, A_m$ et le joueur B, de n stratégies $B_1, B_2, \ldots, B_n$, la matrice $\| a_{ij} \|$ étant donnée.

On demande de trouver la solution du jeu considéré, c'est-à-dire deux stratégies mixtes optimales des joueurs A et B

$$S_A^* = \begin{pmatrix} A_1 & A_2 & \ldots & A_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}; \quad S_B^* = \begin{pmatrix} B_1 & B_2 & \ldots & B_n \\ q_1 & q_2 & \ldots & q_n \end{pmatrix},$$

où $p_1 + \ldots + p_m = 1$; $q_1 + \ldots + q_n = 1$ (certains des nombres p_i et q_j peuvent être nuls).

Notre stratégie optimale S_A^* doit nous garantir un gain au moins égal à ν pour n'importe quelle stratégie adoptée par l'adversaire et égal à ν s'il adopte sa stratégie optimale (ici la stratégie S_B^*). D'une façon analogue, la stratégie S_B^*

doit assurer à l'adversaire une perte ne dépassant pas ν si nous jouons une stratégie quelconque et égale à ν si nous jouons notre stratégie optimale (la stratégie S_A^*).

La valeur du jeu ν n'étant pas connue, admettons qu'elle soit égale à un nombre positif quelconque. Ceci posé nous ne restreignons en aucune manière à la généralité de nos raisonnements; car pour que $\nu > 0$ il suffit évidemment que tous les éléments de la matrice $\| a_{ij} \|$ soient non négatifs. Ce qu'on réalise facilement en ajoutant à ces éléments un nombre L positif suffisamment grand; la valeur du jeu augmente de N sans modifier pour cela la solution du jeu.

Admettons que nous ayons choisi la stratégie optimale S_A^*. Alors notre gain moyen, en supposant que l'adversaire joue la stratégie B_j, sera égal à:

$$a_j = p_1 a_{1j} + p_2 a_{2j} + \ldots + p_m a_{mj}.$$

Notre stratégie optimale S_A^* possède une propriété telle que quel que soit le comportement de l'adversaire elle nous garantit un gain non inférieur à ν; par conséquent, tout nombre a_j ne peut être inférieur à ν. On obtient donc les conditions:

$$\left.\begin{array}{l} p_1 a_{11} + p_2 a_{21} + \ldots + p_m a_{m1} \geqslant \nu, \\ p_1 a_{12} + p_2 a_{22} + \ldots + p_m a_{m2} \geqslant \nu, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p_1 a_{1n} + p_2 a_{2n} + \ldots + p_m a_{mn} \geqslant \nu. \end{array}\right\} \quad (5.1)$$

En divisant les inégalités (5.1) par la grandeur positive ν et en posant

$$\frac{p_1}{\nu} = \xi_1; \quad \frac{p_2}{\nu} = \xi_2; \quad \ldots; \quad \frac{p_m}{\nu} = \xi_m,$$

les conditions (5.1) prennent la forme

$$\left.\begin{array}{l} a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{m1}\xi_m \geqslant 1, \\ a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{m2}\xi_m \geqslant 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{mn}\xi_m \geqslant 1, \end{array}\right\} \quad (5.2)$$

où $\xi_1, \xi_2, \ldots, \xi_m$ sont des nombres non négatifs quelconques. Vu que $p_1 + p_2 + \ldots + p_m = 1$, les grandeurs $\xi_1, \xi_2, \ldots, \xi_m$ satisfont donc à la condition

$$\xi_1 + \xi_2 + \ldots + \xi_m = \frac{1}{\nu}. \tag{5.3}$$

Etant donné que notre tâche est de rendre notre gain garanti le plus grand possible, le second membre de l'égalité (5.3) doit évidemment prendre sa valeur minimale.

Ainsi, le problème posé se ramène au suivant: *trouver des valeurs* $\xi_1, \xi_2, \ldots, \xi_m$ *non négatives satisfaisant aux conditions* (5.2.) *et telles que leur somme*

$$\Phi = \xi_1 + \xi_2 + \ldots + \xi_m$$

soit minimale.

Généralement pour trouver les valeurs extrémales (maximums et minimums) d'une fonction on prend ses dérivées (première et seconde) qu'on égale ensuite à zéro. Mais vu que la fonction Φ que nous devons rendre minimum est linéaire et que ses dérivées par rapport à tous les argumets sont égales à l'unité, c'est-à-dire ne s'annulent nulle part, nous ne pouvons nous servir de cette méthode. Par conséquent, cette fonction atteint son maximum quelque part sur la frontière du domaine de variation des arguments, frontière qui est définie par la non-négativité des arguments et par les conditions (5.2). Il est impossible également de déterminer les valeurs extrémales d'une fonction à l'aide de la dérivation dans les cas où la résolution du jeu implique la connaissance du maximum de la limite inférieure (ou du minimum de la limite supérieure) d'un gain, comme nous l'avons fait en résolvant les jeux de $2 \times n$. En effet, la limite inférieure étant constituée d'une ligne polygonale, le maximum est donc atteint non pas en un point où s'annule la dérivée (un tel point n'existe tout simplement pas), mais sur la frontière d'un tel intervalle ou bien en un point d'intersection des segments de la ligne polygonale.

Ces problèmes qui sont particulièrement importants font l'objet de la *programmation linéaire.*

Celle-ci consiste en ce qui suit.

Soit un système d'équations linéaires:

$$\left.\begin{array}{l} a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{m1}\xi_m = b_1, \\ a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{m2}\xi_m = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{mn}\xi_m = b_n. \end{array}\right\} \quad (5.4)$$

On demande de trouver des $\xi_1, \xi_2, \ldots, \xi_m$ non négatifs, satisfaisant aux conditions (5.4) et rendant minimale la fonction linéaire (ou forme linéaire) homogène des grandeurs $\xi_1, \xi_2, \ldots, \xi_m$

$$\Phi = c_1\xi_1 + c_2\xi_2 + \ldots + c_m\xi_m.$$

Il est aisé de voir que le problème de la théorie des jeux formulé ci-dessus est un cas particulier de celui de la programmation linéaire pour $c_1 = c_2 = \ldots = c_m = 1$.

De prime abord il peut paraître que les conditions (5.2) ne sont pas équivalentes à (5.4) étant donné que dans les premières figurent des inégalités au sens large et dans les secondes, des égalités. Néanmoins, on s'affranchit des signes d'inégalité en introduisant les nouvelles variables fictives non négatives $z_1, z_2, \ldots, z_n$. Les conditions (5.2) s'écrivent alors:

$$\left.\begin{array}{l} a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{m1}\xi_m - z_1 = 1; \\ a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{m2}\xi_m - z_2 = 1; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{mn}\xi_m - z_n = 1. \end{array}\right\} \quad (5.5)$$

La forme Φ que nous devons rendre minimale est alors

$$\Phi = \xi_1 + \xi_2 + \ldots + \xi_m.$$

La méthode de programmation linéaire permet, par un nombre assez restreint d'approximations successives, de

sélectionner les valeurs $\xi_1, \xi_2, \ldots, \xi_m$, satisfaisant aux conditions formulées. Pour plus de clarté nous allons appliquer cette méthode à la résolution de certains jeux concrets.

● EXEMPLE 1. On demande de résoudre le jeu de 3×3 de l'exemple 2 du § 1 représenté par le tableau

A \ B	B_1	B_2	B_3
A_1	2	—3	4
A_2	—3	4	—5
A_3	4	—5	6

Pour rendre tous les a_{ij} non négatifs, ajoutons à tous les éléments de la matrice la grandeur $L = 5$. La matrice prend alors la forme:

A \ B	B_1	B_2	B_3
A_1	7	2	9
A_2	2	9	0
A_3	9	0	11

La valeur du jeu considéré augmente de 5 sans que sa solution en soit modifiée.

Déterminons la stratégie optimale S_A^*. Les conditions (5.2) s'écrivent :

$$\left.\begin{aligned} 7\xi_1 + 2\xi_2 + 9\xi_3 &\geqslant 1,\\ 2\xi_1 + 9\xi_2 \qquad &\geqslant 1,\\ 9\xi_1 + \qquad 11\xi_3 &\geqslant 1,\end{aligned}\right\} \tag{5.6}$$

où $\xi_1 = \frac{p_1}{\gamma}$; $\xi_2 = \frac{p_2}{\gamma}$; $\xi_3 = \frac{p_3}{\gamma}$.

Pour éliminer les signes d'inégalité, introduisons les variables fictives z_1, z_2, z_3, les conditions (5.6) prennent alors la forme :

$$\left.\begin{aligned} 7\xi_1 + 2\xi_2 + 9\xi_3 - z_1 &= 1,\\ 2\xi_1 + 9\xi_2 \qquad - z_2 &= 1,\\ 9\xi_1 + \qquad 11\xi_3 - z_3 &= 1.\end{aligned}\right\} \tag{5.7}$$

La forme linéaire Φ est

$$\Phi = \xi_1 + \xi_2 + \xi_3$$

et doit être rendue minimale.

Si les stratégies de B sont toutes les trois « utiles », les variables fictives z_1, z_2, z_3 s'annulent toutes (c'est-à-dire chaque stratégie B_j de l'adversaire lui assurera un gain égal à la valeur v du jeu). Mais rien *a priori* ne nous permet d'affirmer que toutes les trois stratégies sont « utiles ». Pour nous en assurer, essayons d'exprimer la forme Φ en fonction des variables fictives z_1, z_2, z_3 et voyons si l'on peut la rendre minimale en annulant ces variables. A cet effet résolvons les équations (5.7) par rapport aux variables ξ_1, ξ_2, ξ_3 (c'est-à-dire explicitons ξ_1, ξ_2, ξ_3 en fonction des variables fictives z_1, z_2, z_3) :

$$\left.\begin{aligned} \xi_1 &= \frac{1}{20} - \frac{99}{80} z_1 + \frac{11}{40} z_2 + \frac{81}{80} z_3,\\ \xi_2 &= \frac{1}{10} + \frac{11}{40} z_1 + \frac{1}{20} z_2 - \frac{9}{40} z_3,\\ \xi_3 &= \frac{1}{20} + \frac{81}{80} z_1 - \frac{9}{40} z_2 - \frac{59}{80} z_3.\end{aligned}\right| \tag{5.8}$$

En additionnant ξ_1, ξ_2 et ξ_3, on obtient:

$$\Phi = \frac{1}{5} + \frac{1}{20} z_1 + \frac{1}{10} z_2 + \frac{1}{20} z_3. \quad (5.9)$$

Dans l'équation (5.9) tous les coefficients de z sont positifs; par conséquent, tout accroissement de z_1, z_2,z_3 entraîne un accroissement de la forme Φ. Par conséquent, les valeurs de z_1, z_2, z_3 qui minimisent la forme (5.9) sont

$$z_1 = z_2 = z_3 = 0.$$

Substituant ces valeurs dans l'expression (5.9), on obtient la valeur minimale de la forme Φ:

$$\frac{1}{\nu} = \frac{1}{5},$$

d'où l'on a pour la valeur du jeu

$$\nu = 5.$$

Substituant les valeurs nulles de z_1, z_2, z_3 dans les formules (5.8), on obtient:

$$\xi_1 = \frac{1}{20}; \quad \xi_2 = \frac{1}{10}; \quad \xi_3 = \frac{1}{20},$$

ou, en multipliant ces dernières expressions par ν,

$$p_1 = \frac{1}{4}; \quad p_2 = \frac{1}{2}; \quad p_3 = \frac{1}{4}.$$

Donc la stratégie optimale de A est:

$$S_A^* = \begin{pmatrix} A_1 & A_2 & A_3 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix},$$

c'est-à-dire dans un quart des cas il faut écrire 1, dans la moitié des cas 2 et dans le quart restant 3.

La valeur du jeu $\nu = 5$ étant connue, il est aisé de déterminer la stratégie optimale de l'adversaire

$$S_B^* = \begin{pmatrix} B_1 & B_2 & B_3 \\ q_1 & q_2 & q_3 \end{pmatrix}.$$

Pour la trouver, faisons appel à deux quelconques de nos stratégies « utiles » (A_2 et A_3, par exemple) et écrivons les équations :

$$2q_1 + 9q_2 = 5;$$
$$9q_1 + 11(1 - q_2 - q_1) = 5,$$

d'où $q_1 = q_3 = \frac{1}{4}$; $q_2 = \frac{1}{2}$. La stratégie optimale de l'adversaire coïncide donc avec la nôtre:

$$S_B^* = \begin{pmatrix} B_1 & B_2 & B_3 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}.$$

Revenons maintenant à la forme primitive (non transformée) du jeu considéré. Pour cela il nous faut retrancher de la valeur du jeu $\nu = 5$ le nombre $L = 5$, ajouté aux éléments de la matrice. On obtient donc la valeur du jeu initial $\nu_0 = 0$. Par conséquent, les stratégies optimales assurent aux deux joueurs un gain moyen égal à zéro ; ce jeu est donc indifféremment avantageux ou désavantageux pour les deux adversaires.

● EXEMPLE 2. Un club sportif A peut aligner les équipes A_1, A_2 ou A_3, le club B les équipes B_1, B_2 ou B_3. En s'engageant chacun des deux clubs ignore la composition de l'équipe adverse. Les résultats des rencontres précédentes ont permis d'établir la matrice suivante de probabilité de victoire du club A :

A \ B	B_1	B_2	B_3
A_1	0,8	0,2	0,4
A_2	0,4	0,5	0,6
A_3	0,1	0,7	0,3

On demande de déterminer la fréquence avec laquelle chaque club doit faire varier la composition de son équipe pour s'assurer le plus grand nombre (en moyenne) de victoires.

● SOLUTION. Les valeurs inférieure et supérieure sont 0,4 et 0,6 respectivement. Cherchons la solution du jeu considéré dans le domaine des stratégies mixtes. Pour ne pas faire usage des fractions, multiplions tous les éléments de la matrice ci-dessus par 10 ; la valeur du jeu augmentera ainsi de 10 fois sans que soit modifiée la solution cherchée. On obtient donc la matrice suivante :

A \ B	B_1	B_2	B_3
A_1	8	2	4
A_2	4	5	6
A_3	1	7	3

Les conditions (5.5) prennent alors la forme :

$$\begin{aligned} 8\xi_1 + 4\xi_2 + \xi_3 - z_1 &= 1, \\ 2\xi_1 + 5\xi_2 + 7\xi_3 - z_2 &= 1, \\ 4\xi_1 + 6\xi_2 + 3\xi_3 - z_3 &= 1. \end{aligned} \qquad (5.10)$$

La condition de minimum s'écrit :

$$\Phi = \xi_1 + \xi_2 + \xi_3 = \min.$$

Vérifions maintenant si les stratégies de l'adversaire sont toutes trois « utiles ». Admettons que les variables fictives

z_1, z_2, z_3 soient nulles et résolvons les équations (5.10) par rapport aux variables ξ_1, ξ_2, ξ_3:

$$\left.\begin{aligned}\xi_1 &= \frac{10}{136} + \frac{27}{136} z_1 + \frac{6}{136} z_2 - \frac{23}{136} z_3,\\ \xi_2 &= \frac{12}{136} - \frac{22}{136} z_1 - \frac{20}{136} z_2 + \frac{54}{136} z_3,\\ \xi_3 &= \frac{8}{136} + \frac{8}{136} z_1 + \frac{32}{136} z_2 - \frac{32}{136} z_3,\end{aligned}\right\} \tag{5.11}$$

d'où

$$136\Phi = 30 + 13z_1 + 18z_2 - 51z_3. \tag{5.12}$$

La formule (5.12) montre que l'accroissement des variables z_1 et z_2 entraîne un accroissement de Φ, tandis que l'accroissement de z_3 entraîne la diminution de Φ. Néanmoins, l'accroissement de z_3 doit se faire avec précaution pour que les grandeurs ξ_1, ξ_2, ξ_3 qui sont fonction de z_3 ne deviennent pas négatives. Pour cela annulons z_1 et z_2 dans les seconds membres des égalités (5.11) et augmentons z_3 jusqu'aux limites admissibles (jusqu'à ce que l'une quelconque des grandeurs ξ_1, ξ_2, ξ_3 devienne nulle). De la deuxième égalité (5.11) il résulte que l'augmentation de z_3 n'est pas « dangereuse » pour ξ_2, cette dernière ne peut que s'accroître. En ce qui concerne ξ_1 et ξ_3, on ne peut augmenter z_3 que jusqu'à une certaine limite. En effet ξ_1 s'annule pour $z_3 = \frac{10}{23}$, et ξ_3, pour $z_3 = \frac{1}{4}$. Par conséquent, en donnant à z_3 sa valeur maximale admissible $z_3 = \frac{1}{4}$, nous annulons par là ξ_3.

Pour vérifier si la forme Φ admet un minimum lorsque $z_1 = 0$, $z_2 = 0$, $\xi_3 = 0$, exprimons les autres variables (non nulles) en fonction de z_1, z_2, ξ_3 qui sont nulles par hypothèse.

En résolvant les équations (5.10) par rapport à ξ_1, ξ_2 et z_3, on obtient :

$$\left.\begin{aligned}\xi_1 &= \frac{1}{32}+\frac{5}{32}z_1-\frac{4}{32}z_2+\frac{23}{32}\xi_3,\\ \xi_2 &= \frac{6}{32}-\frac{2}{32}z_1+\frac{8}{32}z_2-\frac{54}{32}\xi_3,\\ \xi_3 &= \frac{8}{32}+\frac{8}{32}z_1+z_2-\frac{136}{32}\xi_3,\end{aligned}\right\}$$

d'où

$$32\Phi = 7 + 3z_1 + 4z_2 + \xi_3. \qquad (5.13)$$

De la formule (5.13) il s'ensuit que tout accroissement des grandeurs z_1, z_2, ξ_3 au-dessus de zéro ne peut qu'accroître la forme Φ. La solution du jeu considéré est donc :

$$z_1 = z_2 = \xi_3 = 0,$$

d'où

$$\xi_1 = \frac{1}{32}; \quad \xi_2 = \frac{3}{16}; \quad z_3 = \frac{1}{4}.$$

Substituant ces valeurs dans la formule (5.13), on détermine la valeur ν du jeu considéré :

$$32\Phi = 7 = \frac{32}{\nu}; \quad \nu = \frac{32}{7}.$$

Notre stratégie optimale est alors :

$$S_A^* = \begin{pmatrix} A_1 & A_2 \\ \frac{1}{7} & \frac{6}{7} \end{pmatrix}.$$

Les stratégies « utiles » (les équipes A_1 et A_2) doivent être adoptées avec les fréquences $\frac{1}{7}$ et $\frac{6}{7}$; A_3 ne doit jamais être utilisée.

Pour trouver la stratégie optimale de l'adversaire, on peut, dans le cas général, procéder comme suit : changer le signe du

gain, ajouter ensuite aux éléments de la matrice, pour les rendre non négatifs, une grandeur constante L et résoudre, enfin, le problème comme nous l'avons fait pour notre stratégie optimale. Cependant la connaissance de la valeur du jeu ν simplifie légèrement ce problème. De plus, dans le cas présent le problème est simplifié davantage par le fait que la solution ne fait intervenir que deux stratégies « utiles » de l'adversaire (B_1 et B_2) car z_3 n'est jamais nulle et par conséquent la stratégie B_3 ne permet pas d'atteindre la valeur du jeu. En choisissant l'une quelconque des stratégies « utiles » du joueur A, A_1 par exemple, on peut déterminer les fréquences q_1 et q_2. Ecrivons pour cela l'équation

$$8q_1 + 2(1 - q_1) = \frac{32}{7},$$

d'où

$$q_1 = \frac{3}{7}; \quad q_2 = \frac{4}{7};$$

la stratégie optimale de l'adversaire est donc:

$$S_B^* = \begin{pmatrix} B_1 & B_2 \\ \frac{3}{7} & \frac{4}{7} \end{pmatrix},$$

ainsi, l'adversaire doit aligner les équipes B_1 et B_2 avec les fréquences $\frac{3}{7}$ et $\frac{4}{7}$ respectivement, et ne jamais solliciter l'équipe B_3.

Revenons à la matrice primitive et déterminons la valeur du jeu

$$\nu_0 = \frac{32}{7} : 10 = 0{,}457.$$

Cela signifie que pour un grand nombre de rencontres le club A remportera 0,457 de tous les matches.

§ 6. MÉTHODES APPROCHÉES DE RÉSOLUTION DES JEUX

Lorsqu'on résout des problèmes pratiques, souvent il n'est pas nécessaire de rechercher la solution exacte: on peut se contenter d'une solution approchée donnant un gain moyen proche de la valeur d'un jeu. Un simple examen de la matrice et la détermination des valeurs inférieure (α) et supérieure (β) du jeu permettent déjà de se faire une idée sur sa valeur v. Dans le cas où α et β sont assez proches il n'est donc pas nécessaire de chercher la solution exacte, il suffit de choisir des stratégies pures minimax. Dans le cas contraire on peut obtenir une solution admissible par des méthodes numériques dont nous allons exposer en bref la méthode *itérative*.

La méthode itérative consiste dans ce qui suit. On fait une « expérience imaginaire » où les joueurs A et B utilisent leurs stratégies. Plus précisément, on a une suite de jeux élémentaires possédant chacun une matrice. Ainsi, nous (le joueur A) commençons par choisir d'une façon arbitraire une de nos stratégies, A_i par exemple. L'adversaire lui oppose aussitôt la stratégie B_j la plus désavantageuse pour nous, c'est-à-dire celle qui minimise notre gain. A la stratégie B_j de l'adversaire, nous opposons une stratégie A_k telle qui maximise notre gain moyen. Ensuite c'est à l'adversaire de jouer. Il nous opposera évidemment une stratégie B_l telle qui minimise notre gain moyen assuré par le couple de stratégies (A_i, A_k) et ainsi de suite. A chaque étape de ce processus itératif chaque joueur oppose à l'adversaire celle de ses stratégies qui est optimale par rapport à tous ses coups précédents qui sont considérés comme une stratégie mixte où les stratégies pures interviennent en proportions correspondant à leur fréquence.

Ce procédé représente en quelque sorte un modèle d'apprentissage où chaque joueur cherche à connaître le comportement de l'autre pour s'opposer le plus efficacement à ses manœuvres.

Si l'on poursuit assez longtemps cette imitation de l'apprentissage, le gain moyen correspondant à un couple de coups (c'est-à-dire à un jeu élémentaire) tendra vers la valeur du jeu et les fréquences $p_1, \ldots, p_m$; $q_1, \ldots, q_n$, des stratégies des joueurs tendront vers les fréquences des stratégies optimales. Les calculs montrent que cette méthode converge lentement, cependant ce n'est pas un obstacle pour les calculateurs.

Illustrons l'application de la méthode itérative d'un jeu de 3×3 résolu dans l'exemple 2 du paragraphe précédent.

Le jeu est donné par la matrice:

A \ B	B_1	B_2	B_3
A_1	8	2	4
A_2	4	5	6
A_3	1	7	3

Dans le tableau 6.1 sont données les 18 premières étapes de l'opération itérative. Dans la première colonne figure le numéro (n) d'un jeu élémentaire (c'est-à-dire de deux coups); dans la deuxième, le numéro (i) de la stratégie choisie par A ; dans les trois colonnes suivantes, le « gain accumulé » au cours des n premiers jeux contre les stratégies B_1, B_2, B_3 de l'adversaire. Le minimum de ces valeurs est souligné d'un trait. Dans la colonne suivante figure le numéro (j) de la stratégie choisie par l'adversaire et dans les trois colonnes suivantes, le gain accumulé au cours des n premiers jeux contre les stratégies A_1, A_2, A_3; le maximum de ces valeurs est surmonté d'un trait. Les valeurs marquées déterminent la stratégie choisie par l'autre joueur. Dans les

Tableau 6.1

n	i	B_1	B_2	B_3	j	A_1	A_2	A_3	$\underline{\nu}$	$\bar{\nu}$	ν^*
1	3	$\underline{1}$	7	3	$\overline{1}$	$\overline{8}$	4	1	1	8	4,50
2	1	9	9	$\underline{7}$	3	$\overline{12}$	10	4	3,50	6,00	4,75
3	1	17	$\underline{11}$	11	2	14	$\overline{15}$	11	3,67	5,00	4,33
4	2	21	$\underline{16}$	17	2	16	$\overline{20}$	18	4,00	5,00	4,50
5	2	25	$\underline{21}$	23	2	18	$\overline{25}$	25	4,20	5,00	4,60
6	2	29	$\underline{26}$	29	2	20	30	$\overline{32}$	4,33	5,33	4,82
7	3	$\underline{30}$	33	32	1	28	$\overline{34}$	33	4,29	4,86	4,57
8	2	$\underline{34}$	38	38	1	36	$\overline{38}$	34	4,25	4,75	4,50
9	2	$\underline{38}$	43	44	1	$\overline{44}$	42	35	4,23	4,89	4,56
10	1	46	$\underline{45}$	48	2	46	$\overline{47}$	42	4,50	4,70	4,60
11	2	$\underline{50}$	50	54	1	$\overline{54}$	51	43	4,55	4,91	4,72
12	1	58	$\underline{52}$	58	2	56	$\overline{56}$	50	4,33	4,66	4,49
13	2	62	$\underline{57}$	64	2	58	$\overline{61}$	57	4,38	4,70	4,54
14	2	66	$\underline{62}$	70	2	60	$\overline{66}$	64	4,43	4,71	4,56
15	2	70	$\underline{67}$	76	2	62	71	$\overline{71}$	4,47	4,73	4,60
16	3	$\underline{71}$	74	79	1	70	$\overline{75}$	72	4,44	4,69	4,56
17	2	$\underline{75}$	79	85	1	78	79	73	4,41	4,65	4,53
18	2	$\underline{79}$	84	91	1	$\overline{86}$	83	74	4,39	4,78	4,58
...	...	...	...	...	...	...	...	...	...	...	...

colonnes suivantes sont respectivement inscrits : le gain moyen minimal $\underline{\nu}$ qui est égal au gain minimal accumulé divisé par le nombre n de jeux ; le gain moyen maximal $\overline{\nu}$ égal au gain maximal accumulé divisé par n et, enfin, leur moyenne arithmétique $\nu^* = \frac{\underline{\nu}+\overline{\nu}}{2}$. Lorsque n croît, les trois grandeurs $\underline{\nu}$, $\overline{\nu}$ et ν^* tendront vers la valeur du jeu ν, ν^* plus rapidement que les deux autres.

Cet exemple montre que la méthode itérative converge assez lentement, cependant, si petit soit-il, ce calcul permet de trouver la valeur approchée du jeu et fait apparaître la prédominance des stratégies « utiles ». Les calculateurs améliorent nettement le rendement de cette méthode.

L'avantage de la méthode itérative est que les calculs à effectuer ne se compliquent que légèrement avec l'augmentation des nombres m et n de stratégies.

§ 7. MÉTHODES DE RÉSOLUTION DE CERTAINS JEUX INFINIS

On appelle jeu infini un jeu où l'un au moins des participants dispose d'un nombre illimité de stratégies. Les méthodes générales de résolution de tels jeux sont loin d'être définitivement élaborées. Néanmoins certains cas particuliers où la solution est assez simple, méritent d'être étudiés vu leur importance pratique.

Examinons un jeu où chacun des deux participants A et B dispose d'une infinité de stratégies ; pour A ces stratégies correspondent aux différentes valeurs d'un paramètre continu x, pour B à celles de y. Dans le cas considéré, le jeu est déterminé non plus par la matrice $\| a_{ij} \|$, mais par une certaine fonction $a(x, y)$ de deux variables continues que nous appellerons dans la suite *fonction de gain* (il importe de remarquer que la fonction $a(x, y)$ n'est pas forcément continue). La fonction de gain $a(x, y)$ peut être représentée par

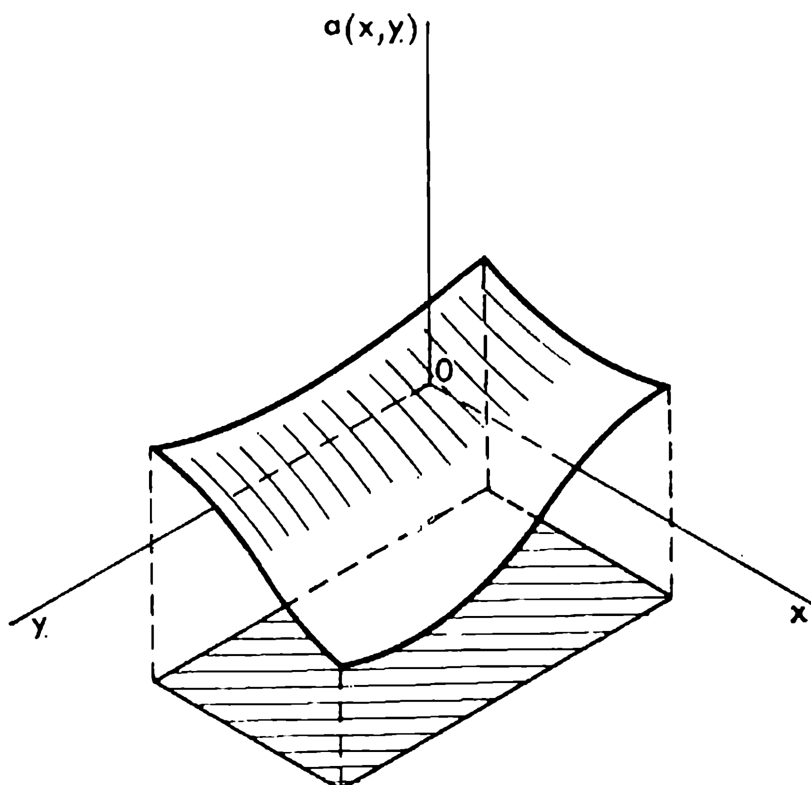


Fig. 16

une certaine surface géométrique $a(x, y)$ au-dessus du domaine de variation des arguments (x, y) (fig. 16).

L'examen de la fonction de gain $a(x, y)$ est analogue à celui d'une matrice de paiement. On commence par déterminer la valeur inférieure α du jeu ; pour cela on détermine pour chaque x le minimum de la fonction $a(x, y)$ par rapport à tous les y:

$$\min_{y} a(x, y);$$

et, ensuite, le maximum de ces valeurs par rapport à tous les x (maximin):

$$\alpha = \max_{x} \min_{y} a(x, y).$$

La valeur supérieure du jeu considéré (minimax) se détermine de façon analogue:

$$\beta = \min_{y} \max_{x} a(x, y).$$

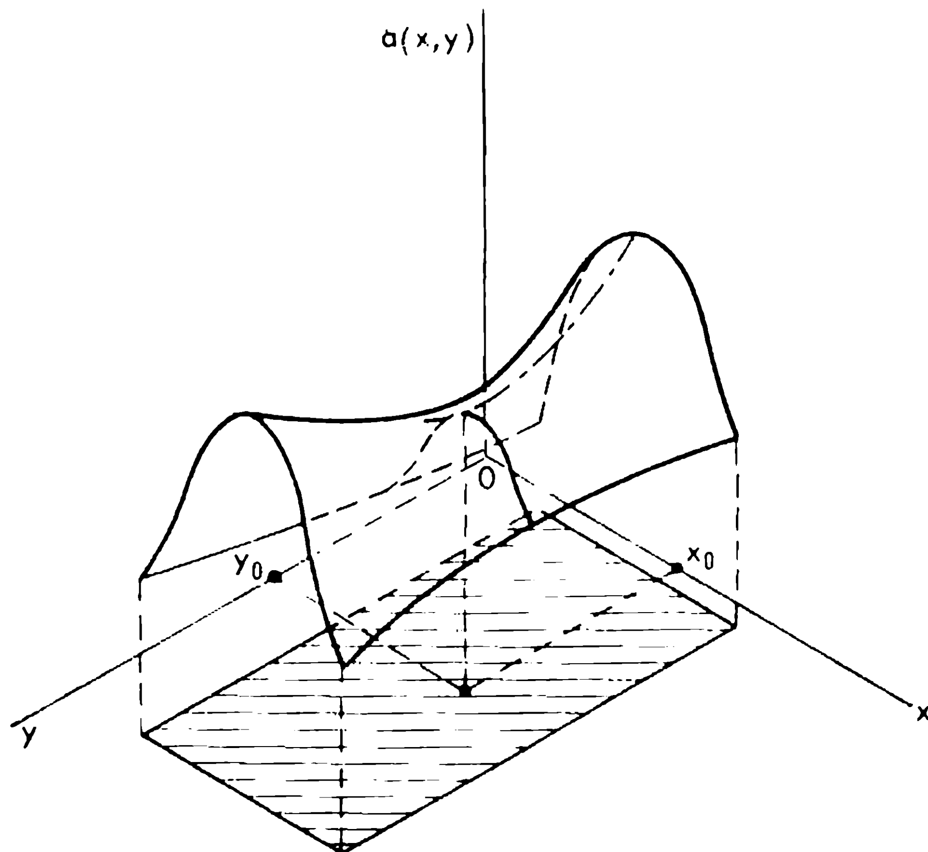


Fig. 17

Examinons maintenant le cas où $\alpha = \beta$. Etant donné que la valeur du jeu ν est toujours comprise entre α et β, leur valeur générale est précisément ν.

L'égalité $\alpha = \beta$ signifie que la surface $a(x, y)$ présente un *point selle*, c'est-à-dire un point de coordonnées x_0, y_0 tel que $a(x, y)$ est simultanément minimale par rapport à y et maximale par rapport à x (fig. 17).

La valeur de $a(x, y)$ en ce point est précisément la valeur du jeu ν:

$$\nu = a(x_0, y_0).$$

L'existence du point selle signifie que le jeu infini considéré admet une solution dans le domaine des stratégies pures; x_0, y_0 représentent donc les stratégies pures optimales de A et B. Dans le cas général, lorsque $\alpha \neq \beta$, le jeu peut

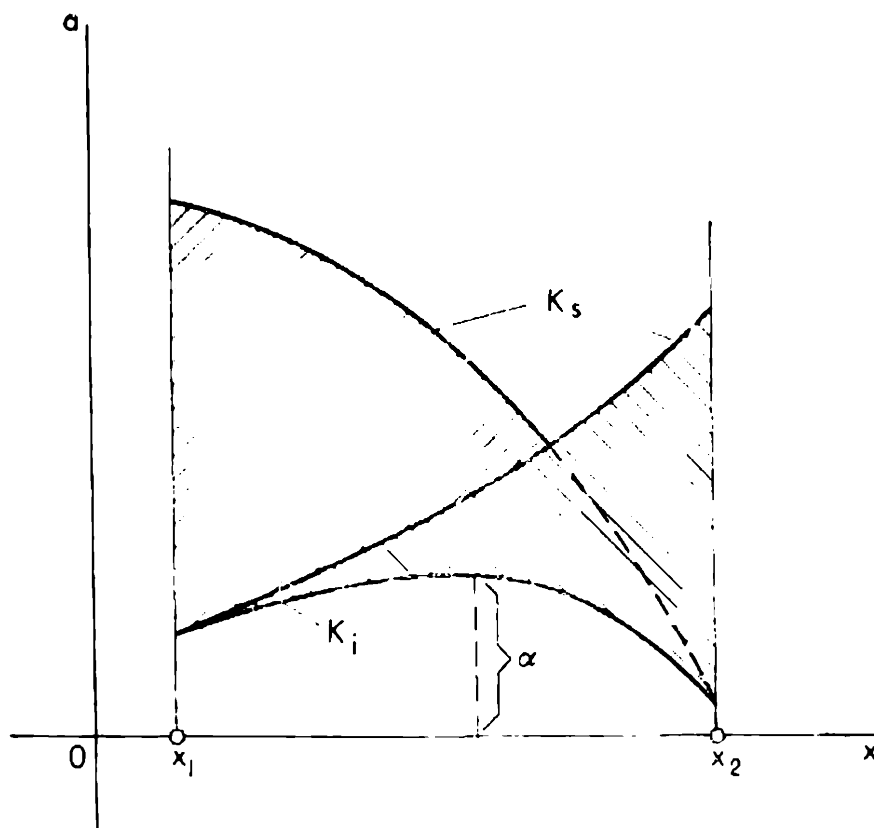


Fig. 18

n'avoir de solution que dans le domaine des stratégies mixtes (pas forcément une solution unique). Dans les jeux infinis la stratégie mixte représente une certaine distribution des probabilités pour les stratégies x et y qui sont considérées comme des valeurs aléatoires. Cette distribution peut être continue et déterminée par les densités $f_1(x)$ et $f_2(y)$; ou bien discrète, et alors les stratégies optimales sont constituées de stratégies pures choisies avec des probabilités quelconques non nulles.

Dans le cas où un jeu infini n'a pas de point selle, ses valeurs inférieure et supérieure admettent une interprétation géométrique fort concrète. Considérons un jeu infini de fonction de gain $a(x, y)$ et dont les stratégies x, y couvrent continûment les segments (x_1, x_2) et (y_1, y_2) des axes. Pour déterminer la valeur inférieure α du jeu, il faut « regarder » la surface $a(x, y)$ à partir de l'axe y, c'est-à-dire la projeter

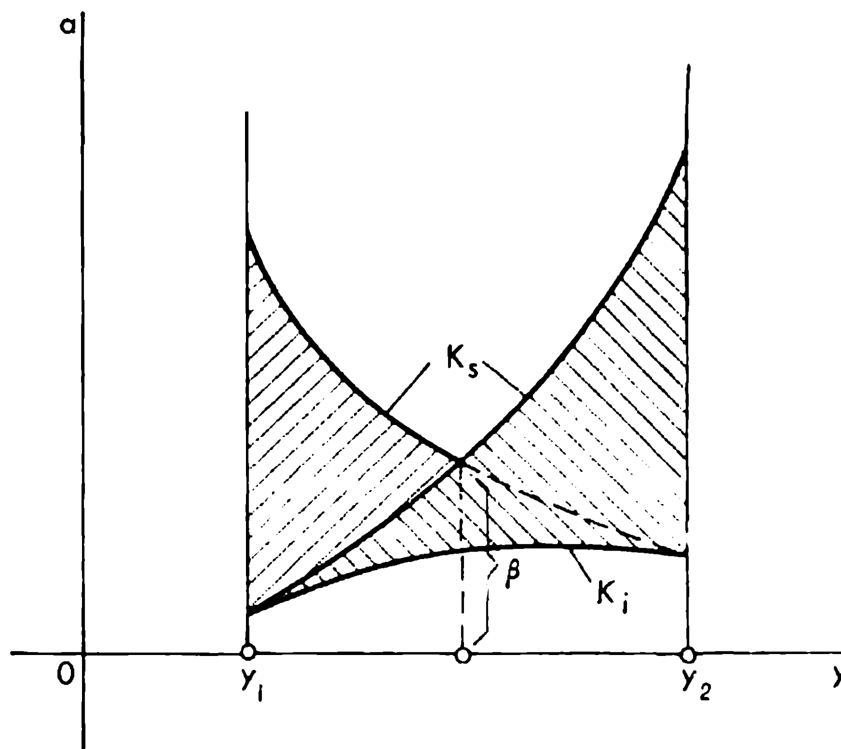


Fig. 19

sur le plan xOa (fig. 18). On obtient ainsi une certaine figure limitée à gauche et à droite par les droites $x = x_1$ et $x = x_2$, d'en haut et d'en bas par les courbes K_s et K_i. La valeur inférieure α du jeu n'est évidemment rien d'autre que l'ordonnée maximale de la courbe K_i. De façon analogue, pour déterminer la valeur supérieure β du jeu il faut « regarder » la surface $a(x, y)$ à partir de l'axe Ox (la projeter sur le plan yOa) et trouver l'ordonnée minimale de la limite supérieure K_s de la projection obtenue (fig. 19).

Considérons maintenant deux exemples fort élémentaires de jeux infinis.

● EXEMPLE 1. Deux joueurs A et B disposent chacun d'un nombre illimité de stratégies possibles x et y telles que $0 \leqslant x \leqslant 1$; $0 \leqslant y \leqslant 1$.

La fonction de gain est donnée par l'expression

$$a(x, y) = (x - y)^2.$$

Trouver la solution de ce jeu.

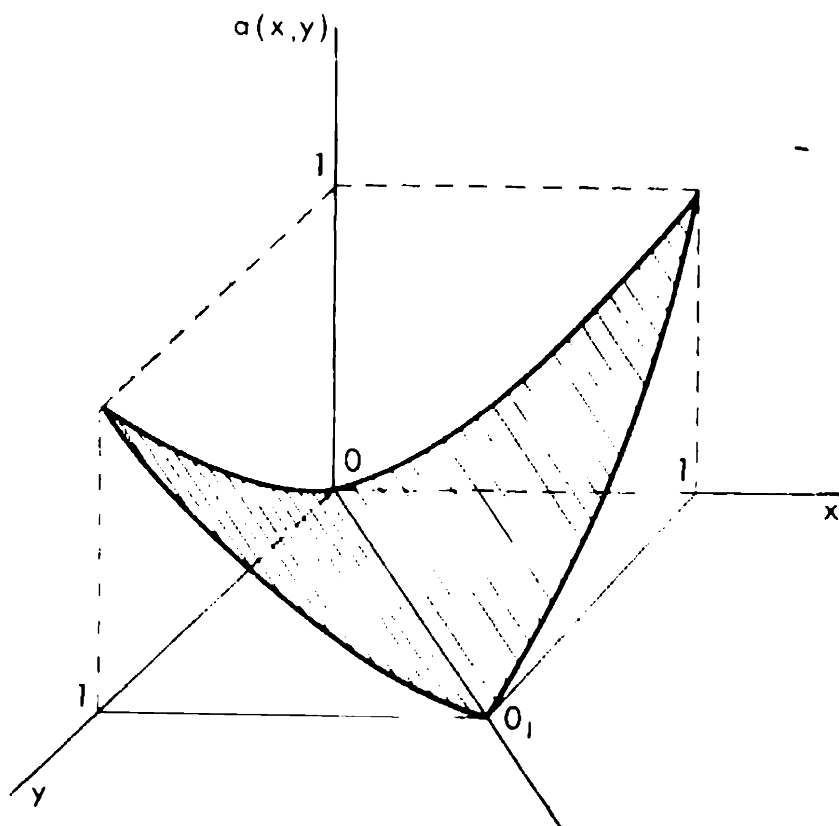


Fig. 20

● SOLUTION. La surface $a(x, y)$ représente un cylindre parabolique (fig. 20) et n'a pas de point selle. Déterminons la valeur inférieure du jeu ; il est évident que $\min_y a(x, y) = 0$ pour tous les x ; d'où

$$\alpha = \max_x \min_y a(x, y) = 0.$$

Trouvons maintenant la valeur supérieure de ce jeu. On a pour y fixé

$$\max_x (x - y)^2.$$

Dans le cas considéré, le maximum s'obtient toujours sur l'une des bornes de l'intervalle considéré (pour $x = 0$ ou $x = 1$), c'est-à-dire il est égal à la plus grande des valeurs y^2 et $(1 - y)^2$. Traçons les graphiques de ces fonctions

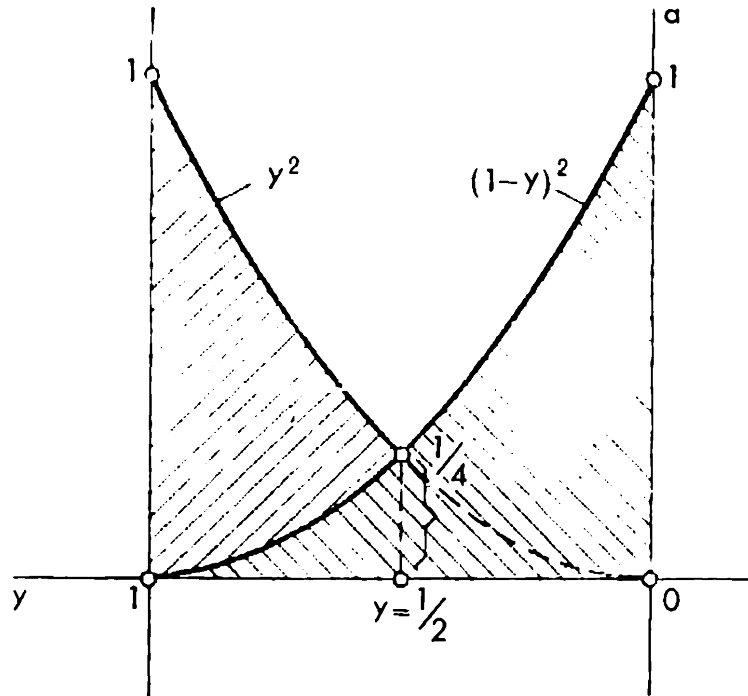


Fig. 21

(fig. 21), autrement dit la projection de la surface $a(x, y)$ sur le plan yOa. Sur la fig. 21 le graphique de la fonction $\max\limits_{x} (x-y)^2$ est tracé en gras. Il est aisé de voir que sa valeur minimale est atteinte pour $y = \frac{1}{2}$ et égale à $\frac{1}{4}$. Par conséquent, la valeur supérieure du jeu considéré est $\beta = \frac{1}{4}$.

Dans le cas examiné, la valeur supérieure du jeu coïncide avec la valeur du jeu ν. En effet, le joueur A peut adopter la stratégie mixte $S_A = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ où les valeurs extrémales $x=0$ et $x=1$ alternent avec les mêmes fréquences ; donc, pour toute stratégie y du joueur B le gain moyen de A sera égal à:

$$\frac{1}{2} y^2 + \frac{1}{2} (1-y)^2.$$

On voit sans peine que quel que soit y compris entre 0 et 1, cette grandeur ne sera pas inférieure à $\frac{1}{4}$:

$$\frac{1}{2} y^2 + \frac{1}{2} (1-y)^2 \geqslant \frac{1}{4}.$$

De cette façon, en adoptant la stratégie mixte mentionnée le joueur A peut s'assurer un gain égal à la valeur supérieure du jeu ; étant donné que la valeur du jeu ne peut pas dépasser la valeur supérieure, la stratégie S_A est donc optimale

$$S_A = S_A^*.$$

Il ne nous reste qu'à déterminer la stratégie optimale du joueur B.

Il est évident que si la valeur ν du jeu coïncide avec sa valeur supérieure β, la stratégie optimale du joueur B se confondra toujours avec sa stratégie pure minimax lui assurant la valeur supérieure du jeu. Dans le cas considéré cette stratégie optimale est $y_0 = \frac{1}{2}$. En effet, si B adopte cette stratégie, le gain du joueur A, quelle que soit sa manière d'agir, n'excédera jamais $\frac{1}{4}$. Cette affirmation découle de l'inégalité évidente

$$\left(x - \frac{1}{2}\right)^2 = x(x-1) + \frac{1}{4} \leqslant \frac{1}{4}.$$

● EXEMPLE 2. La partie A (« nous ») tire sur un avion de B. Pour échapper à notre tir, l'ennemi doit manœuvrer avec une certaine surcharge y dont les valeurs peuvent varier de $y = 0$ (mouvement rectiligne) à $y = y_{\max}$ (vol suivant un cercle de courbure maximale). Prenons $y_{\max}$ pour unité de mesure, c'est-à-dire posons $y_{\max} = 1$.

Dans notre lutte contre l'adversaire, nous pouvons utiliser des appareils de pointage dont le principe de fonctionnement est fondé sur telle ou telle hypothèse relative au

déplacement de la cible pendant le vol de l'obus. Dans ces conditions la surcharge x peut prendre une valeur quelconque comprise entre 0 et 1.

Notre objectif est donc d'abattre l'avion ennemi, celui de l'adversaire, d'échapper à notre feu. La probabilité d'être atteint s'exprime, pour x et y donnés, par la formule approchée suivante

$$a(x, y) = pe^{-k(x-y)^2},$$

où y est la surcharge choisie par l'adversaire; x celle prise en compte par l'appareil de pointage.

On demande de déterminer les stratégies optimales des deux adversaires.

● SOLUTION. Il est aisé de voir que si l'on pose $p = 1$ on ne modifie en rien la solution du jeu. La fonction de gain $a(x, y)$ prend alors la forme de la surface représentée sur la fig. 22. C'est une surface cylindrique dont les génératrices sont parallèles à la bissectrice de l'angle xOy formé par les axes x et y; en coupant cette surface par un plan perpendiculaire à l'une des génératrices on obtient une courbe ayant la forme d'une courbe de distribution normale.

En utilisant l'interprétation géométrique des valeurs inférieure et supérieure du jeu exposée plus haut, on trouve que $\beta = 1$ (fig. 23) et $\alpha = e^{-\frac{k}{4}}$ (fig. 24).

Le jeu considéré n'a pas de point selle; sa solution est donc comprise dans le domaine des stratégies mixtes. Ce problème est donc en quelque sorte analogue à celui de l'exemple précédent. En effet, pour k petit la fonction $e^{-k(x-y)^2}$ se comporte à peu près comme la fonction $-(x - y)^2$ et la solution du jeu se déduit de la solution du problème précédent, en permutant A et B, c'est-à-dire notre stratégie optimale se confondra avec la stratégie pure $x = \frac{1}{2}$

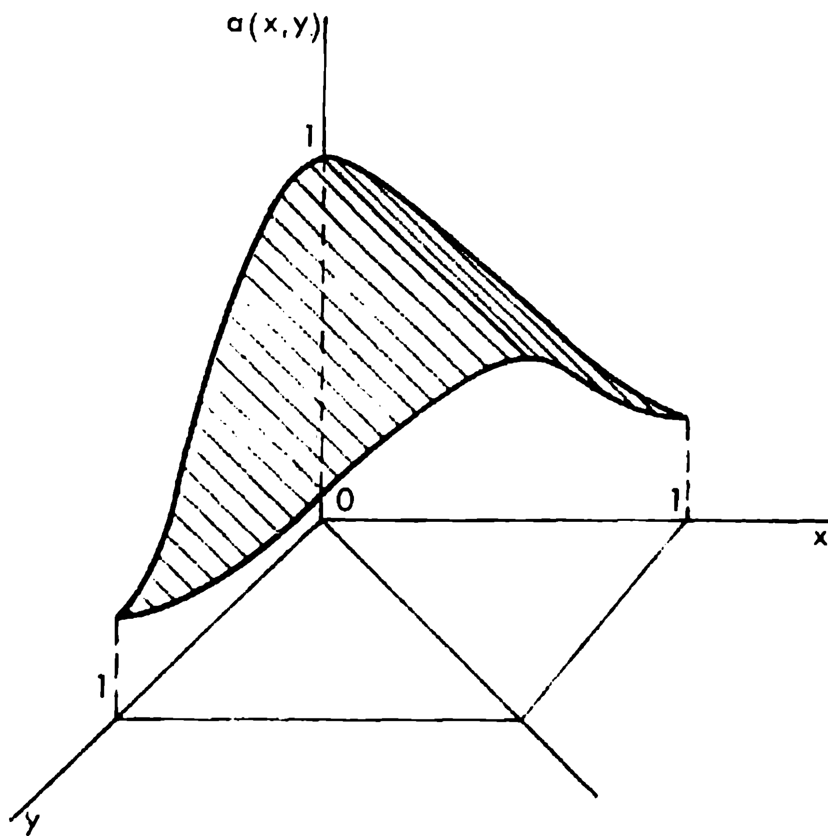


Fig. 22

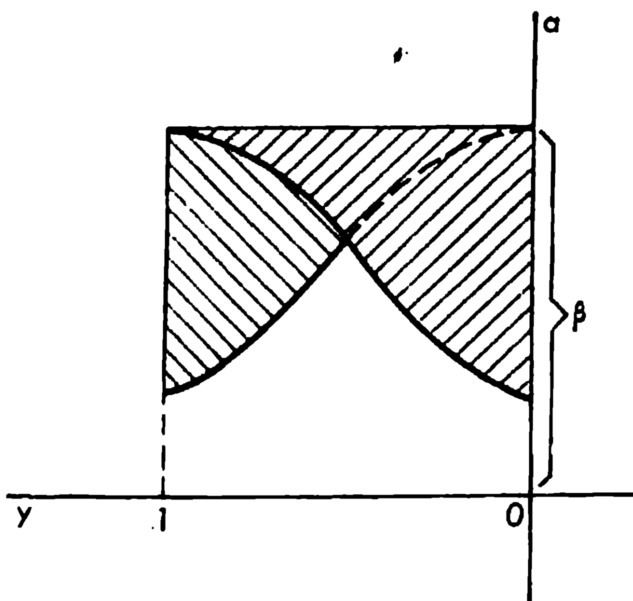


Fig. 23

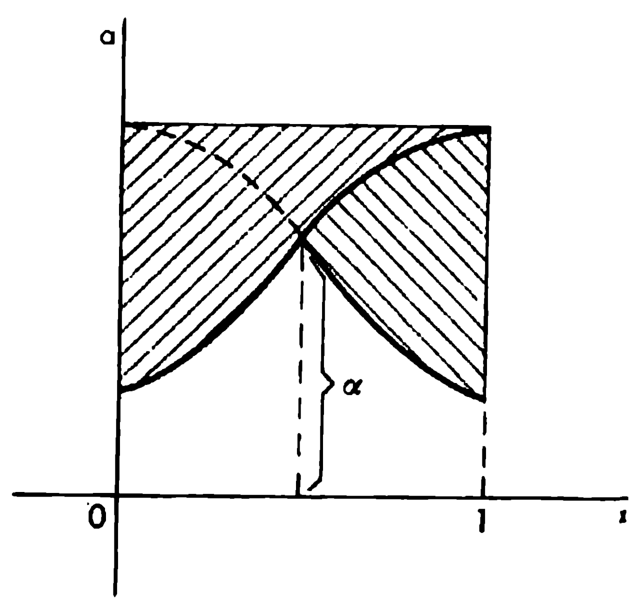


Fig. 24

et la stratégie optimale de l'adversaire

$$S_B^* = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

consistera à utiliser avec les mêmes fréquences les stratégies extrémales $y = 0$ et $y = 1$. Cela signifie que dans tous les cas notre appareil de pointage doit être fixé pour la surcharge $x = \frac{1}{2}$ tandis que l'adversaire doit, dans une moitié des cas, éviter toute manœuvre pour recourir dans l'autre à des manœuvres maximales.

Il est facile de démontrer que cette solution sera valable pour $k \leqslant 2$. En effet, si l'adversaire adopte la stratégie

$$S_B = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

et nous, la stratégie x, le gain moyen s'exprime par la fonction

$$a(x) = \frac{1}{2}\left(e^{-kx^2} + e^{-k(1-x)^2}\right),$$

qui lorsque $k \leqslant 2$, admet un maximum égal à la valeur inférieure α du jeu pour $x = \frac{1}{2}$. Par conséquent, la stratégie S_B assure à l'adversaire une perte n'excédant pas α, d'où il résulte que la valeur inférieure du jeu α coïncide avec la valeur du jeu v. Pour $k > 2$, la fonction $a(x)$ admet deux maximums en x_0 et $1 - x_0$ symétriques par rapport à $x = \frac{1}{2}$ (fig. 25) et de plus x_0 dépend de k.

Il est clair que lorsque $k = 2$ on a $x_0 = 1 - x_0 = \frac{1}{2}$; si k augmente, les points x_0 et $1 - x_0$ s'écartent et tendent vers les points extrêmes (0 et 1). Par conséquent, la solution du jeu dépendra de k. Posons $k = 3$, par exemple, et cherchons la solution du jeu ; à cet effet déterminons l'abs-

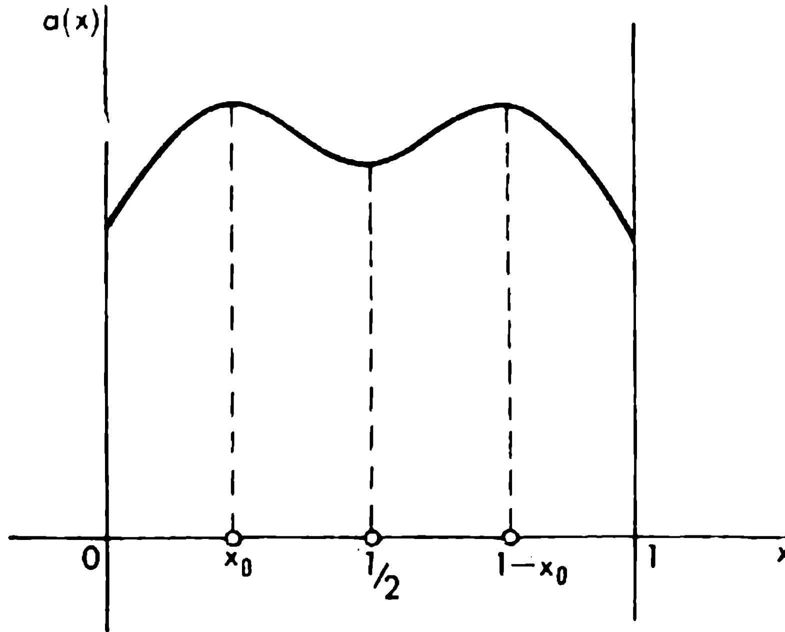


Fig. 25

cisse x_0 du maximum de la courbe $a(x)$. En égalant la dérivée de la fonction $a(x)$ à zéro, on tire x_0 de l'équation

$$xe^{-3x^2} = (1-x)\,e^{-3(1-x)^2}.$$

Cette équation a trois racines: $x = \frac{1}{2}$ (où la fonction atteint son minimum), x_0 et $1 - x_0$ (où la fonction atteint ses maximums). En résolvant numériquement cette équation, on obtient

$$x_0 \approx 0{,}07\,;$$
$$1 - x_0 \approx 0{,}93.$$

Démontrons maintenant que dans le cas présent la solution du jeu est constituée par le couple suivant de stratégies:

$$S_A^* = \begin{pmatrix} x_0 & 1-x_0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad S_B^* = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Le gain moyen assuré par notre stratégie S_A^* et la stratégie y de l'adversaire sera alors

$$a_1(y) = \frac{1}{2}\left(e^{-3(0{,}07-y)^2} + e^{-3(0{,}93-y)^2}\right).$$

Déterminons maintenant le minimum $a_1(y)$ lorsque $0 < y < 1$. Etant symétrique par rapport à $y = \frac{1}{2}$, la fonction $a_1(y)$ ne peut avoir qu'un seul ou deux maximums; elle atteint donc son minimum soit au milieu de l'intervalle (0, 1), soit à ses extrémités. Ainsi, en posant $y = 0$ (où $y = 1$), on a

$$a_1(0) = a_1(1) = \frac{1}{2}(e^{-3\cdot 0,07^2} + e^{-3\cdot 0,93^2}) = 0,530.$$

En y posant $y = \frac{1}{2}$, on trouve

$$a_1\left(\frac{1}{2}\right) = e^{-3\cdot 0,43^2} = 0,574,$$

c'est-à-dire une valeur supérieure à $a_1(0)$; par conséquent, la valeur du jeu n'est pas inférieure à $a_1(0)$:

$$\nu \geqslant \frac{1}{2}(e^{-3x_0^2} + e^{-3(1-x_0)^2}) = 0,530.$$

Supposons maintenant que l'adversaire utilise la stratégie S_B^* et nous la stratégie x. Le gain moyen sera alors

$$a_2(x) = \frac{1}{2}(e^{-3x^2} + e^{-3(1-x)^2}). \qquad (7.2)$$

Nous avons choisi x_0 de façon que l'expression (7.2) soit maximisée pour $x = x_0$; par conséquent,

$$a_2(x) \leqslant \frac{1}{2}(e^{-3x_0^2} + e^{-3(1-x_0)^2}) = 0,530,$$

c'est-à-dire la stratégie S_B^* permet à l'adversaire de ne pas avoir une perte dépassant 0,530; donc la valeur $\nu = 0,530$ est justement la valeur du jeu et les stratégies S_A^* et S_B^* constituent la solution du problème. En d'autres termes, nous devons utiliser avec la même fréquence les appareils de pointage pour $x = 0,07$ et $x = 0,93$, l'adversaire, quant à lui, doit, avec la même fréquence soit éviter la manœuvre, soit manœuvrer avec une surcharge maximale.

Remarquons que le gain $\nu = 0{,}530$ dépasse notablement la valeur inférieure du jeu

$$\alpha = e^{-\frac{k}{4}} = e^{-0,75} = 0{,}472,$$

que nous pourrions nous assurer en adoptant la stratégie maximin $x_0 = \frac{1}{2}$.

Une méthode pratique de résolution des jeux infinis est celle qui consiste à les ramener approximativement à des jeux finis. Dans ce cas toute une série de stratégies possibles de chaque joueur se trouvent réunies en une seule stratégie. Un tel procédé ne fournit évidemment qu'une solution approchée du jeu, mais dans la plupart des cas pratiques la solution exacte n'est pas indispensable.

Néanmoins il importe de souligner que ce procédé peut faire apparaître des solutions dans le domaine des stratégies mixtes lorsque la solution du jeu infini peut être obtenue avec des stratégies pures, c'est-à-dire lorsque ce jeu infini possède un point selle. Si en ramenant un jeu infini à un jeu fini, on obtient une solution mixte composée seulement de deux stratégies « utiles » voisines, il vaut la peine d'essayer d'utiliser une stratégie pure comprise entre ces dernières.

Notons pour conclure qu'à la différence des jeux finis les jeux infinis peuvent ne pas avoir de solution. En voici un exemple. Deux joueurs citent chacun un nombre entier quelconque. Celui qui a cité le plus grand nombre gagne un rouble. Si les deux joueurs citent le même nombre la partie est nulle. Il est aisé de voir que ce jeu n'admet pas de solution. Il existe cependant des classes de jeux infinis admettant une solution *a priori*. On peut démontrer en particulier que si dans un jeu infini les stratégies possibles x, y des joueurs A, B remplissent continûment certains intervalles et que la fonction de gain $a(x, y)$ soit continue, alors ce jeu admet toujours une solution (avec les stratégies pures ou les stratégies mixtes).

A. Solodovnikov

SYSTÈMES D'INÉGALITÉS LINÉAIRES

INTRODUCTION

Les inégalités du premier degré ou, comme on les appelle autrement, *inégalités linéaires* sont de la forme

$$ax + by + c \geqslant 0$$

(pour plus de simplicité on s'est borné à donner ici une inégalité à deux inconnues x et y). La théorie des systèmes d'inégalités linéaires est un chapitre assez simple mais fort intéressant des mathématiques, car, géométriquement parlant, un système d'inégalités linéaires à deux ou trois inconnues définit soit un domaine convexe polygonal dans un plan soit, respectivement, un corps convexe polyèdre dans un espace. Ainsi, par exemple, la théorie des polyèdres convexes, branche de la géométrie vieille comme le monde, devient une des parties de la théorie des systèmes d'inégalités linéaires. Mais, d'autre part, cette même théorie possède certains chapitres si chers à tout algébriste, comme, par exemple, l'analogie remarquable entre les propriétés des systèmes d'inégalités linéaires et celles des systèmes d'*équations linéaires* (ces derniers étant d'ailleurs étudiés il y a longtemps et bien en détail).

Tout récemment encore on croyait que les inégalités linéaires resteront toujours l'objet d'études purement mathématiques. Mais au beau milieu des années quarante, après l'apparition de la *programmation linéaire*, branche toute nouvelle des mathématiques appliquées très importante pour l'économie et la technique, la situation a radicalement changé. La programmation linéaire n'est au fond que l'une des branches (bien que fort importante) de la théorie des systèmes d'inégalités linéaires.

Ce petit ouvrage se propose comme but de familiariser le lecteur aux divers aspects de la théorie des systèmes d'inégalités linéaires tels que l'interprétation géométrique du problème et les méthodes de résolution de ces systèmes



intimement liées à celle-ci, certaines propriétés purement algébriques des systèmes linéaires, les problèmes fondamentaux de la programmation linéaire. La lecture de cet ouvrage n'impose aucunes connaissances spéciales dépassant le programme scolaire.

Arrêtons-nous brièvement sur l'histoire de la théorie des systèmes d'inégalités linéaires. Bien que d'après son essence même cette théorie doive, paraît-il, se rapporter à des branches les plus fondamentales et les plus élémentaires des mathématiques, elle n'a pas attiré, jusqu'à ces derniers temps, l'attention des mathématiciens.

A partir des dernières années du siècle écoulé, certains ouvrages consacrés aux différentes propriétés des systèmes d'inégalités linéaires faisaient leur apparition. A ce propos on peut citer les noms des mathématiciens ayant beaucoup contribué à l'étude de ces systèmes tels que H. Minkowski (un des plus grands géomètres du siècle dernier auquel on attribue surtout de remarquables travaux sur les ensembles convexes et un cours de géométrie), G. Voronoï (un des fondateurs de l'« école de Pétersbourg de la théorie des nombres »), A. Haar (mathématicien hongrois devenu célèbre grâce à ses travaux sur l'intégration des groupes), H. Weil (un des plus célèbres mathématiciens de la première moitié de notre siècle).

Il est à souligner que la théorie des systèmes d'inégalités linéaires n'a trouvé son développement qu'à partir des années 40-50 de notre siècle quand les progrès rapides des disciplines appliquées (les programmations linéaire, convexe et d'autres formes de la « programmation mathématique », comme, par exemple, la « théorie des jeux », etc.) ont rendu absolument nécessaire l'étude systématique et approfondie des inégalités linéaires. Aujourd'hui la liste plus ou moins complète de tous les ouvrages et articles consacrés à l'étude des inégalités linéaires comporterait des centaines de titres.

§ 1. QUELQUES FAITS EMPRUNTÉS À LA GÉOMÉTRIE ANALYTIQUE

A. *Opérations sur des points.* Introduisons un système de coordonnées rectangulaires dans un plan. Le fait qu'un point M a dans ce système les coordonnées x et y peut être inscrit comme suit :

$$M = (x, y) \text{ ou } M (x, y) \text{ tout simplement.}$$

Le système de coordonnées permet d'effectuer sur les points d'un plan certaines opérations, à savoir : *l'opération d'addition des points et celle de multiplication d'un point par un nombre.*

La première de ces opérations se détermine comme suit : si $M_1 = (x_1, y_1)$ et $M_2 = (x_2, y_2)$, alors

$$M_1 + M_2 = (x_1 + x_2, y_1 + y_2).$$

Ainsi, l'addition des points se ramène à l'addition de leur coordonnées identiques.

Le sens géométrique de cette opération est également fort simple (fig. 1) : le point $M_1 + M_2$ représente donc le quatrième sommet du parallélogramme construit sur les segments OM_1 et OM_2 (O étant l'origine des coordonnées). Les trois autres sommets de ce parallélogramme sont les points M_1, O, M_2.

Ce même fait peut d'ailleurs être exprimé autrement : le point $M_1 + M_2$ s'obtient du point M_2 par une translation parallèle en direction du segment OM_1 à la distance égale à la longueur de ce dernier.

La multiplication du point $M (x, y)$ par un nombre arbitraire k s'opère suivant la règle que voici :

$$kM = (kx, ky).$$

Le sens géométrique de cette opération est encore plus simple que dans le cas précédent : le point $M' = kM$ se trouve, pour $k > 0$, sur le rayon OM et $OM' = k \cdot OM$; pour $k < 0$

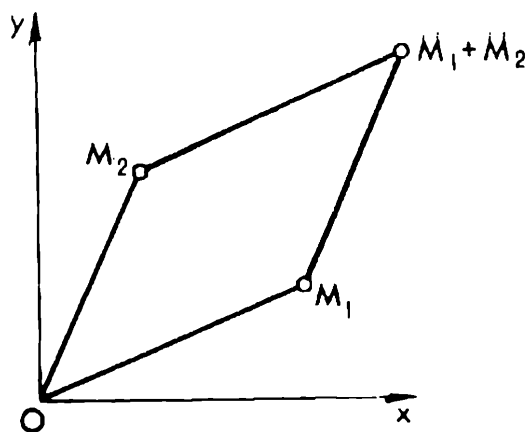


Fig. 1

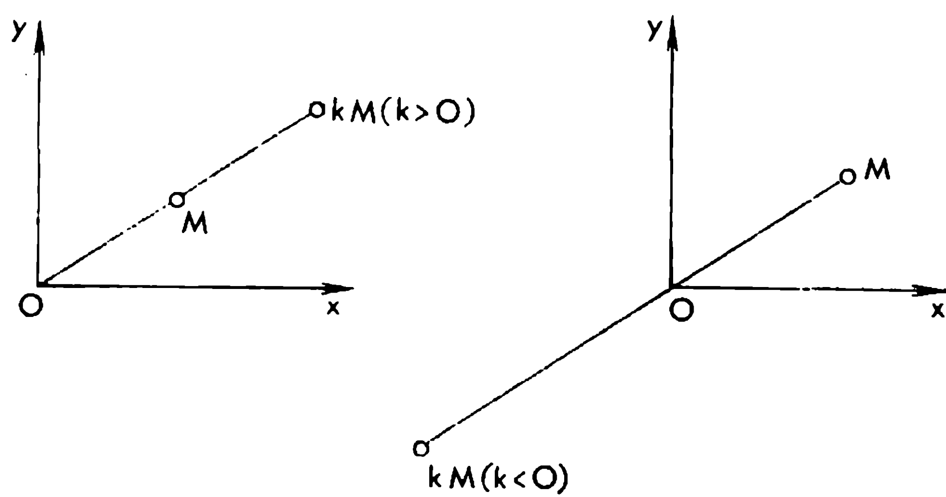


Fig. 2

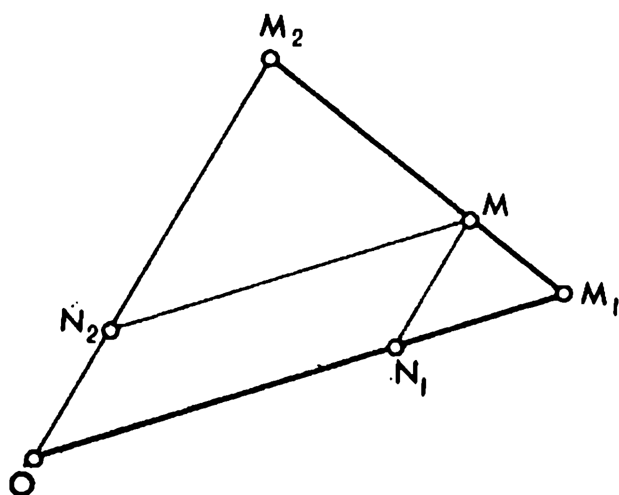


Fig. 3

ce même point M' se trouve sur le prolongement du rayon OM allant au-delà du point O, et $OM' = |k| \cdot OM$ (fig. 2).

La démonstration de l'interprétation géométrique de ces deux opérations servira au lecteur d'un bon exercice *).

Les opérations ainsi introduites permettent de traduire bien facilement les faits géométriques en langage algébrique. En voici quelques exemples.

1) *Le segment M_1M_2 se compose de tous les points de la forme*

$$s_1M_1 + s_2M_2,$$

où s_1, s_2 sont les deux nombres non négatifs quelconques dont la somme vaut 1.

Dans ce cas le fait purement géométrique, l'appartenance d'un point au segment M_1M_2 s'exprime sous la forme d'une relation algébrique $M = s_1M_1 + s_2M_2$, avec les restrictions mentionnées plus haut imposées à s_1, s_2.

Pour la démonstration, considérons un point arbitraire M appartenant au segment M_1M_2. En traçant par le point M des droites parallèles à OM_2 et OM_1, on obtient un point N_1 sur le segment OM_1 et un point N_2 sur le segment OM_2 (fig. 3). Soient

$$s_1 = \frac{M_2M}{M_2M_1}, \quad s_2 = \frac{M_1M}{M_1M_2};$$

les nombres s_1 et s_2 étant non négatifs dont la somme vaut 1. En raison de la similitude des triangles correspondants, on obtient

$$\frac{ON_1}{OM_1} = \frac{M_2M}{M_2M_1} = s_1, \quad \frac{ON_2}{OM_2} = \frac{M_1M}{M_1M_2} = s_2,$$

---

*) Si au moins le lecteur ignore les fondements de la théorie des vecteurs. Du point de vue vectoriel nos opérations signifient que le point $M_1 + M_2$ est l'extrémité du vecteur $\overrightarrow{OM_1} + \overrightarrow{OM_2}$ et le point kM est l'extrémité du vecteur $k \cdot \overrightarrow{OM}$ (à condition, bien sûr, que son origine coïncide avec le point O).

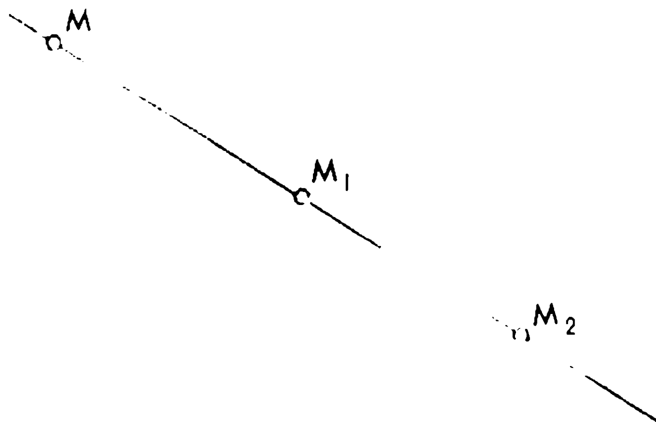


Fig. 4

d'où il s'ensuit $N_1 = s_1M_1$, $N_2 = s_2M_2$. Mais $M = N_1 + N_2$, par conséquent, $M = s_1M_1 + s_2M_2$. Il est à remarquer enfin que lorsque le point M parcourt le segment M_1M_2 dans la direction de M_1 à M_2, le nombre s_2 parcourt toutes les valeurs de 0 à 1. La proposition 1) est donc démontrée.

2) *Tout point M de la droite M_1M_2 peut être mis sous la forme*

$$tM_1 + (1 - t)\,M_2,$$

où t est un point quelconque.

En effet, si le point M appartient au segment M_1M_2, notre assertion résulte de ce qui a été démontré plus haut. Supposons maintenant que le point M n'appartienne pas au segment M_1M_2. Les deux cas suivants peuvent alors se présenter : ou bien le point M_1 appartient au segment MM_2 (voir fig. 4) ou bien le point M_2 se trouve sur le segment MM_1. Supposons, par exemple, qu'ait lieu le premier cas. Alors, d'après ce qu'on a démontré plus haut, on a

$$M_1 = sM + (1 - s)\,M_2 \quad (0 < s < 1),$$

d'où il s'ensuit que

$$M = \frac{1}{s}\,M_1 - \frac{1-s}{s}\,M_2 = tM_1 + (1 - t)\,M_2,$$

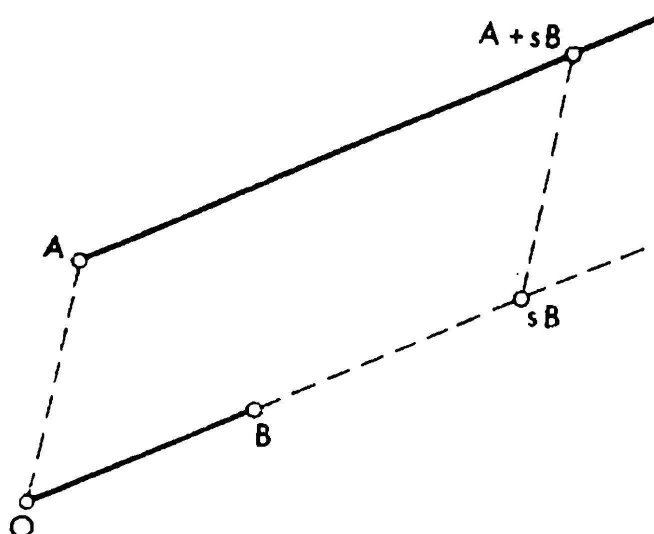


Fig. 5

où $t = \frac{1}{s}$. Nous laissons au lecteur le soin d'examiner le cas où le point M_2 appartient au segment MM_1.

3) *Si le paramètre s augmente de* 0 *à* ∞, *le point sB parcourt le rayon OB* *), *et le point* $A + sB$ *parcourt alors le rayon issu du point A dans le sens OB. Si s diminue de* 0 *à* $-\infty$, *les points sB et* $A + sB$ *parcourent les rayons complémentaires par rapport à ceux mentionnés plus haut.*

De la proposition 3) il s'ensuit que lorsque s varie de $-\infty$ à $+\infty$ le point $A + sB$ parcourt la droite passant par A et parallèle à OB.

Les opérations d'addition et de multiplication des points par un nombre peuvent évidemment s'effectuer sur les points dans un espace. Dans ce cas-là on a par définition,

$$M_1 + M_2 = (x_1 + x_2,\ y_1 + y_2,\ z_1 + z_2),$$
$$kM = (kx,\ ky,\ kz).$$

Il est clair que toutes les propositions démontrées plus haut resteront valides pour un espace.

Pour conclure ce petit chapitre, adoptons une convention qui nous permettra par la suite de formuler certains

*) Le point B est supposé différent de l'origine des coordonnées O.

faits d'une façon plus claire et concise. Notamment si $\mathscr{K}$ et $\mathscr{L}$ sont deux ensembles des points quelconques (dans un plan ou dans un espace), convenons d'entendre par la suite par leur « somme » $\mathscr{K} + \mathscr{L}$ l'ensemble de tous les points de la forme $K + L$, où K est un point arbitraire appartenant à $\mathscr{K}$ et L un point arbitraire appartenant à $\mathscr{L}$.

En mathématiques il existe un système de notations spéciales pour désigner l'appartenance d'un point quelconque à un ensemble déterminé. Notamment pour montrer que le point M appartient à l'ensemble $\mathscr{M}$, on écrit $M \in \mathscr{M}$ (le symbole $\in$ signifie « est élément de »). $\mathscr{K} + \mathscr{L}$ est un ensemble de tous les points de la forme $K + L$, où $K \in \mathscr{K}$ et $L \in \mathscr{L}$.

En partant de l'interprétation géométrique de l'addition des points, il est aisé d'énoncer la règle d'addition des ensembles des points $\mathscr{K}$ et $\mathscr{L}$. Cette règle consiste en ceci : il faut construire pour chaque point $K \in \mathscr{K}$ un ensemble qui s'obtient de $\mathscr{L}$ par une translation parallèle à la longueur du segment OK pour réunir ensuite tous les ensembles ainsi obtenus en un seul qui sera exactement $\mathscr{K} + \mathscr{L}$.

Citons maintenant quelques exemples.

1. Soient un ensemble $\mathscr{K}$ composé d'un seul point K et $\mathscr{L}$ un ensemble arbitraire des points. L'ensemble $K + \mathscr{L}$ provient d'une translation parallèle de l'ensemble $\mathscr{L}$ à la longueur du segment OK (fig. 7). En particulier, si $\mathscr{L}$ est une droite, $K + \mathscr{L}$ est également une droite parallèle à $\mathscr{L}$. Si, de plus, la droite $\mathscr{L}$ passe par l'origine des coordonnées, $K + \mathscr{L}$ est une droite parallèle à $\mathscr{L}$ et passant par le point K (fig. 8).

2. Soient deux segments $\mathscr{K}$ et $\mathscr{L}$ (dans un plan ou dans un espace) non parallèles entre eux (fig. 9). L'ensemble $\mathscr{K} + \mathscr{L}$ est alors un parallélogramme dont les côtés sont égaux et parallèles à $\mathscr{K}$ et $\mathscr{L}$ (respectivement). Quelle figure obtiendra-t-on pour les segments $\mathscr{K}$ et $\mathscr{L}$ parallèles?

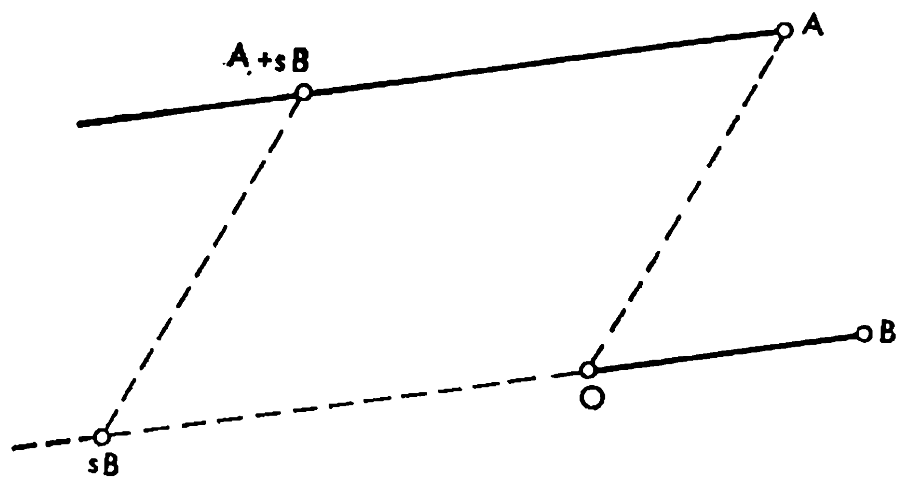


Fig. 6

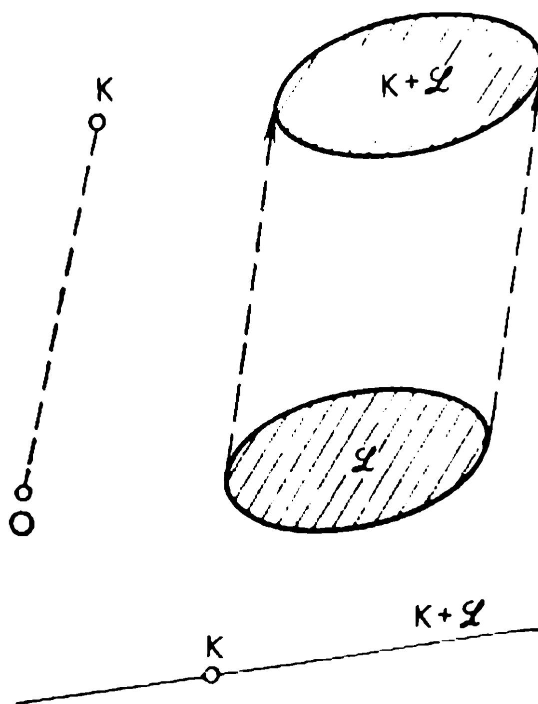


Fig. 7

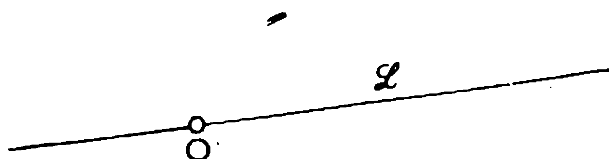


Fig. 8

3. Soient $\mathscr{K}$ un plan et $\mathscr{L}$ un segment quelconque non parallèle à $\mathscr{K}$. L'ensemble $\mathscr{K} + \mathscr{L}$ représente dans ce cas une partie de l'espace comprise entre deux plans parallèles à $\mathscr{K}$ (fig. 10).

4. Soient $\mathscr{K}$ et $\mathscr{L}$ deux cercles de rayons r_1 et r_2 et centrés aux points P_1 et P_2 (respectivement) situés dans un même plan π. L'ensemble $\mathscr{K} + \mathscr{L}$ est alors un cercle de rayon $r_1 + r_2$ centré au point $P_1 + P_2$ et situé dans un plan parallèle à π (fig. 11).

B. *Le sens géométrique d'une équation et d'une inégalité du premier degré à deux ou trois inconnues.* Examinons une équation du premier degré à deux inconnues x et y:

$$ax + by + c = 0. \tag{1}$$

En considérant x et y comme les coordonnées d'un point dans un plan, il est naturel de poser la question: qu'est-ce qu'on a pour ensemble qui est formé dans un plan par les points dont les coordonnées satisfont à l'équation (1), ou plus brièvement, quel est l'ensemble des points défini par l'équation (1)?

Voici la réponse bien que le lecteur la connaisse déjà: *l'ensemble des points défini par l'équation (1) est une droite dans un plan.* En effet, si $b \neq 0$, l'équation (1) prend la forme

$$y = kx + p,$$

mais cette dernière définit, comme on le sait, une droite. Si $b = 0$, cette équation prend alors la forme

$$x = h$$

et définit une droite parallèle à l'axe des ordonnées.

Une question analogue se pose naturellement lorsqu'il s'agit de l'inégalité

$$ax + by + c \geqslant 0. \tag{2}$$

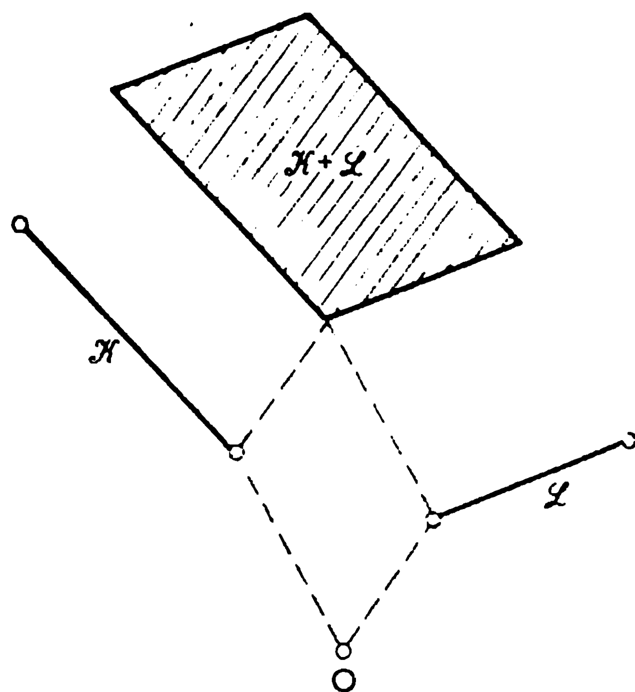


Fig. 9

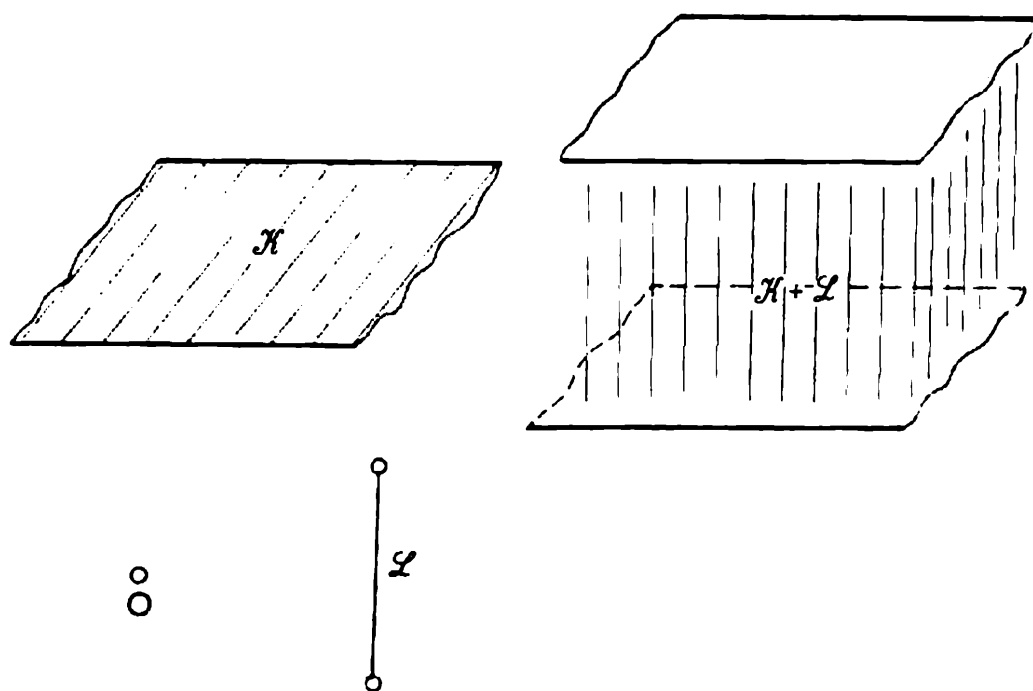


Fig. 10

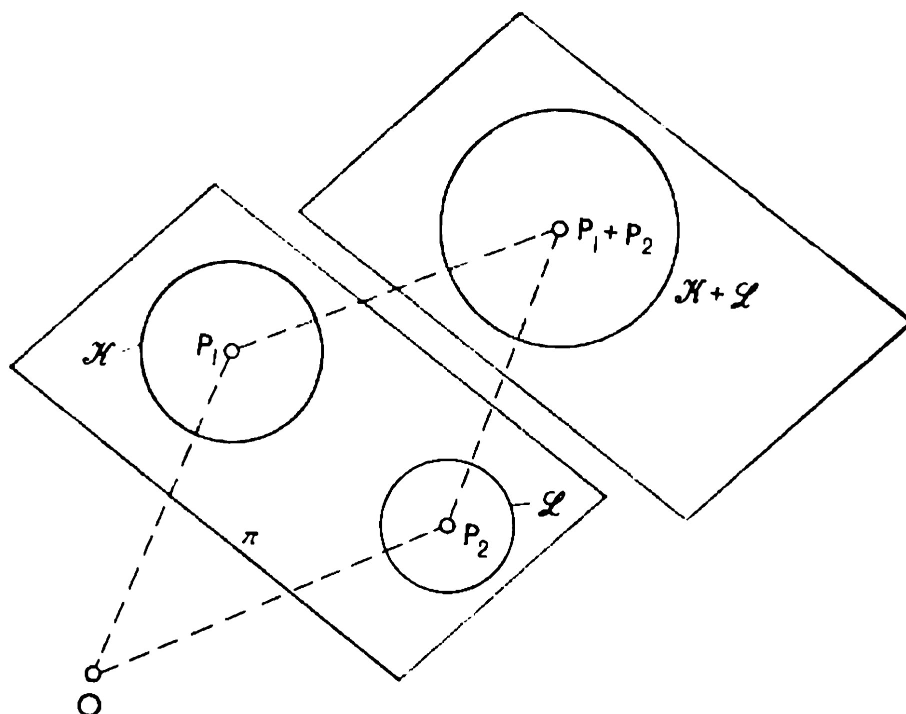


Fig. 11

Quel est l'ensemble des points dans un plan qui est défini par l'inégalité (2)?

Dans ce cas aussi la réponse est simple. Si $b \neq 0$, l'inégalité donnée se ramène à l'une des formes

$$y \geqslant kx + p \text{ ou } y \leqslant kx + p.$$

Il est aisé de voir que la première de ces inégalités est vérifiée par tous les points situés « au-dessus » de la droite $y = kx + p$ ou appartenant à cette droite même tandis que la deuxième est vérifiée par tous les points situés « au-dessous » de cette droite ou appartenant à cette droite même (fig. 12). Si $b = 0$, l'inégalité (2) prend l'une des formes suivantes

$$x \geqslant h \quad \text{ou} \quad x \leqslant h;$$

la première est vérifiée par tous les points situés « à droite » de la droite $x = h$ ou appartenant à cette dernière, la seconde est vérifiée par tous les points situés « à gauche » de la droite $x = h$ ou appartenant à cette dernière (fig. 13).

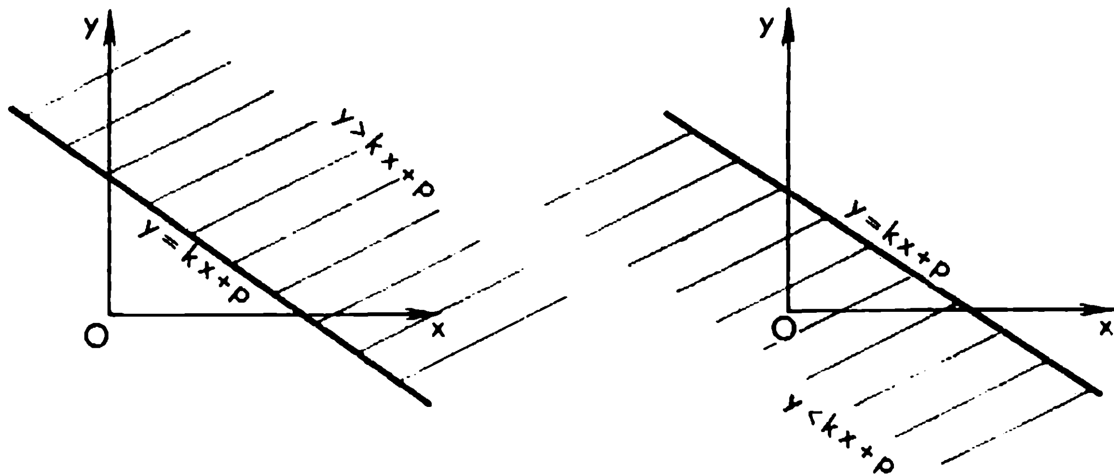


Fig. 12

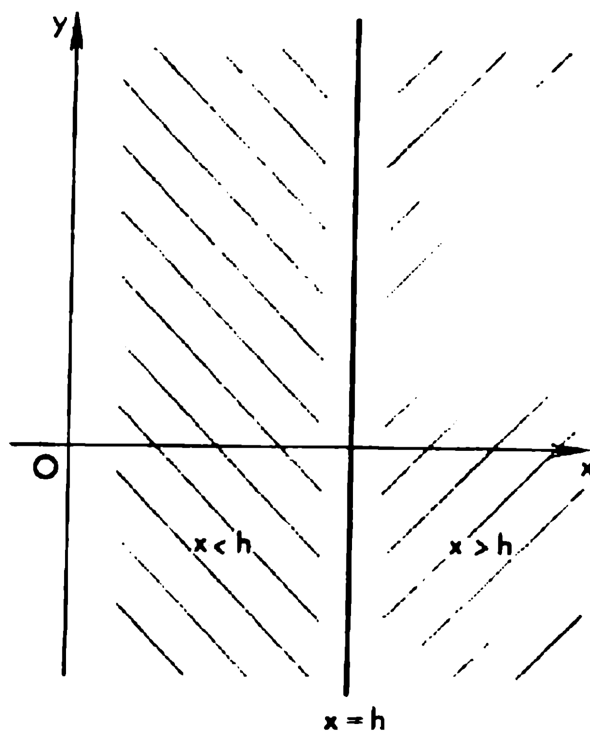


Fig. 13

Ainsi, *l'équation* (*1*) *définit dans un plan de coordonnées une droite et l'inégalité* (*2*) *l'un des deux demi-plans en lesquels cette droite divise ce plan* (la droite même est considérée comme appartenant à n'importe lequel de ces deux demi-plans).

Nous passons maintenant à la résolution des problèmes analogues par rapport à l'équation

$$ax + by + cz + d = 0 \tag{3}$$

et à l'inégalité

$$ax + by + cz + d \geqslant 0; \tag{4}$$

tout en considérant, évidemment, x, y, z comme des coordonnées d'un point quelconque *dans l'espace*. Il est aisé de voir que l'on obtient le résultat suivant.

● THÉORÈME. *L'équation* (3) *définit dans l'espace un certain plan et l'inégalité* (4) *l'un des deux demi-espaces en lesquels ce plan divise tout l'espace* (le plan lui-même est considéré comme appartenant à n'importe lequel des deux demi-espaces).

● DÉMONSTRATION. L'un au moins de trois nombres a, b et c n'est pas nul ; soit, par exemple, $c \neq 0$. L'équation (3) prend alors la forme:

$$z = kx + ly + p. \tag{5}$$

Désignons par $\mathcal{L}$ l'ensemble de tous les points M (x, y, z) vérifiant l'équation (5). Il nous faut démontrer que $\mathcal{L}$ est un plan.

Déterminons d'abord les points de $\mathcal{L}$ appartenant au plan de coordonnées yOz. A cet effet, il faut poser dans (5) $x = 0$. On obtient

$$z = ly + p. \tag{6}$$

Ainsi, l'intersection de $\mathcal{L}$ avec le plan yOz est une droite u définie dans ce plan par l'équation (6) (fig. 14).

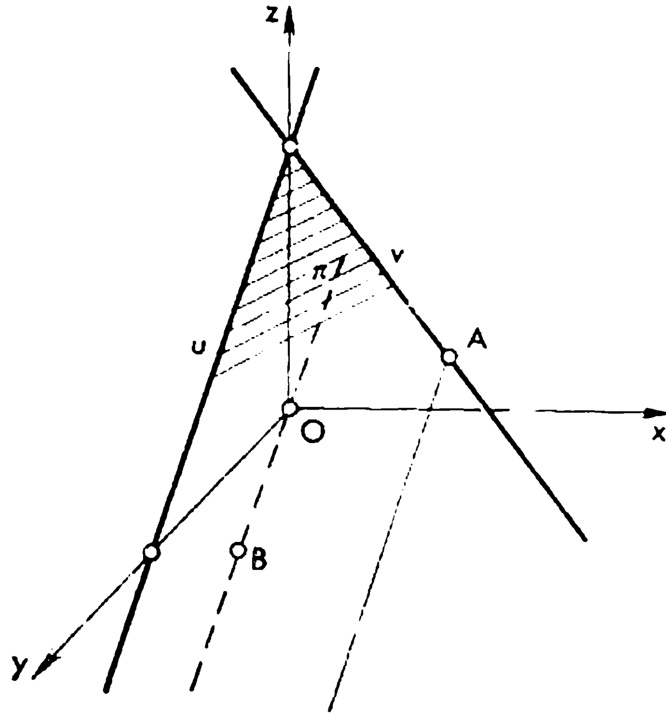


Fig. 14

De façon analogue on trouve que l'intersection de $\mathcal{L}$ avec le plan xOz est une droite v définie dans ce plan par l'équation

$$z = kx + p. \tag{7}$$

Les deux droites u et v passent par le point P (0, 0, p).

Désignons par π un plan contenant les droites u et v. Nous allons démontrer que π *appartient* à l'ensemble $\mathcal{L}$.

Pour cela il suffit d'établir le fait suivant: une droite passant par tout point $A \in v$ parallèlement à u appartient à $\mathcal{L}$.

Choisissons d'abord un point B quelconque tel que $OB \parallel u$. Dans le plan yOz l'équation $z = ly + p$ définit la droite u; donc, l'équation $z = ly$ définit une droite parallèle à u et passant par l'origine des coordonnées (en pointillé sur la fig. 14). On peut prendre pour B un point de coordonnées $y = 1$, $z = l$ situé sur cette droite.

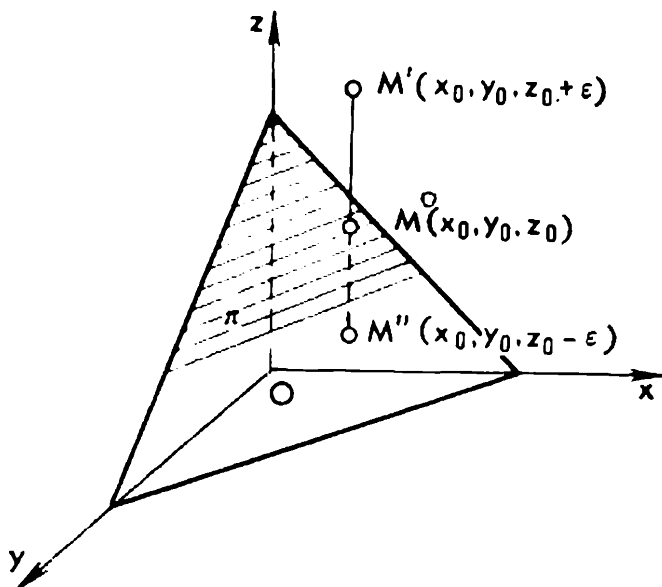


Fig. 15

Un point arbitraire $A \in v$ a pour ses coordonnées x, 0, $kx + p$. Les coordonnées du point B choisi sont 0, 1, l. Une droite passant par A parallèlement à u se compose de points

$$A + sB = (x,\ 0,\ kx + p) + s\,(0,\ 1,\ l) =$$
$$= (x,\ s,\ kx + p + sl),$$

où s est un nombre arbitraire (voir la proposition 3) de A). Il est aisé de voir que les coordonnées du point $A + sB$ vérifient l'équation (5), c'est-à-dire que $A + sB \in \mathscr{L}$. Ce fait montre que le plan π appartient entièrement à l'ensemble $\mathscr{L}$.

Il nous reste à démontrer que l'ensemble $\mathscr{L}$ *coïncide* avec π, c'est-à-dire que cet ensemble ne possède aucun point en dehors de π.

Considérons pour cela trois points : le point $M\ (x_0,\ y_0,\ z_0)$ se trouvant dans le plan π ; le point $M'\ (x_0,\ y_0,\ z_0 + \varepsilon)$ situé « au-dessus » du plan π ($\varepsilon > 0$) et, enfin, le point $M''\ (x_0,\ y_0,\ z_0 - \varepsilon)$ situé « au-dessous » du plan π (fig. 15). Vu que $M \in \pi$, on a $z_0 = kx_0 + ly_0 + p$ et, par consé-

quent,

$$z_0 + \varepsilon > kx_0 + ly_0 + p, \; z_0 - \varepsilon < kx_0 + ly_0 + p.$$

D'où il s'ensuit que les coordonnées du point M' vérifient l'inégalité stricte suivante

$$z > kx + ly + p$$

et celles du point M'' l'inégalité stricte

$$z < kx + ly + p.$$

Cela signifie que les points M' et M'' n'appartiennent pas à l'ensemble $\mathscr{L}$. Ceci démontre que $\mathscr{L}$ se confond avec le plan π. Nos raisonnements nous amènent de plus à conclure que l'ensemble de tous les points vérifiant l'inégalité

$$ax + by + cz + d \geqslant 0$$

est l'un des deux demi-espaces (« supérieur » ou « inférieur ») en lesquels le plan π partage l'espace entier.

§ 2. SENS GÉOMÉTRIQUE D'UN SYSTÈME D'INÉGALITÉS LINÉAIRES À DEUX OU TROIS INCONNUES

Soit

$$\left.\begin{array}{l} a_1x + b_1y + c_1 \geqslant 0, \\ a_2x + b_2y + c_2 \geqslant 0, \\ \dots\dots\dots\dots \\ a_mx + b_my + c_m \geqslant 0 \end{array}\right\} \qquad (1)$$

un système d'inégalités à deux inconnues x et y.

La première inégalité de ce système définit dans le plan de coordonnées xOy un certain demi-plan Π_1, la deuxième — un demi-plan Π_2, etc. Si un couple quelconque de nombres x, y vérifie toutes les inégalités (1), le point cor-

respondant M (x, y) appartient alors à tous les demi-plans Π_1 Π_2, . . ., Π_m simultanément. En d'autres termes, le point M appartient à *l'intersection* (à la partie commune) des demi-plans considérés. Il est aisé de voir que l'intersection d'un nombre fini de demi-plans est un certain domaine polygonal $\mathcal{K}$. La figure 16 montre l'un des domaines possibles de $\mathcal{K}$. La partie intérieure du contour de ce domaine est marquée de hachures pour montrer de quel côté de la droite donnée est situé le demi-plan correspondant. Les flèches y montrent la même chose.

Le domaine $\mathcal{K}$ est appelé *domaine de solutions du système* (1). Notons tout de suite que le domaine de solutions n'est pas nécessairement limité; ainsi, l'intersection de plusieurs demi-plans peut aboutir à l'apparition d'un demi-plan illimité comme, par exemple, celui représenté sur la figure 17. En tenant compte du fait que la frontière du domaine $\mathcal{K}$ se compose de segments de droite (ou de droites entières), on dit que $\mathcal{K}$ est un *domaine polygonal* (il est à remarquer que dans le cas où le domaine $\mathcal{K}$ est limité on l'appelle tout court *polygone* *) *de solutions du système* (1)). Naturellement le cas est possible où il n'y aura aucun point appartenant simultanément à tous les demi-plans considérés, c'est-à-dire lorsque le domaine $\mathcal{K}$ sera « vide », cela signifie que le système (1) est contradictoire. Le cas en question est représenté sur la figure 18.

Le domaine de solutions $\mathcal{K}$ est toujours convexe. Rappelons que selon la définition générale, *l'ensemble des points* (*dans un plan ou dans un espace*) *est dit convexe si avec ses*

---

*) Pour éviter tout malentendu, il convient de faire ici une remarque. Dans les programmes scolaires de géométrie, on entend par « polygone » une *ligne* fermée composée de segments de droites. Néanmoins dans les ouvrages traitant les problèmes d'inégalités linéaires ce terme désigne non pas la ligne elle-même mais *l'ensemble de tous les points d'un plan qu'elle renferme* (c'est-à-dire situés à l'intérieur ou sur la ligne elle-même). Dans ce qui suit nous allons comprendre le terme « polygone » dans ce dernier sens.

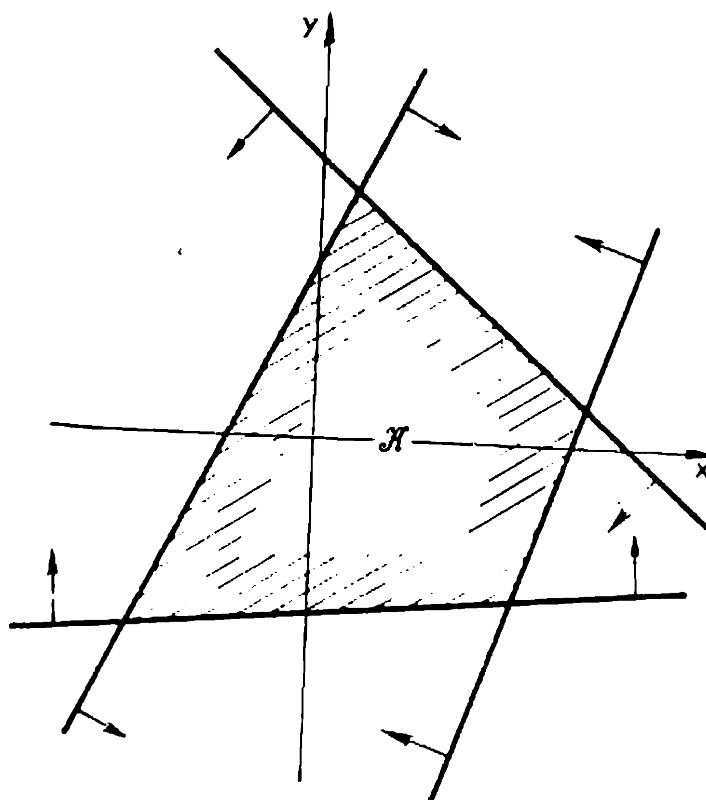


Fig. 16

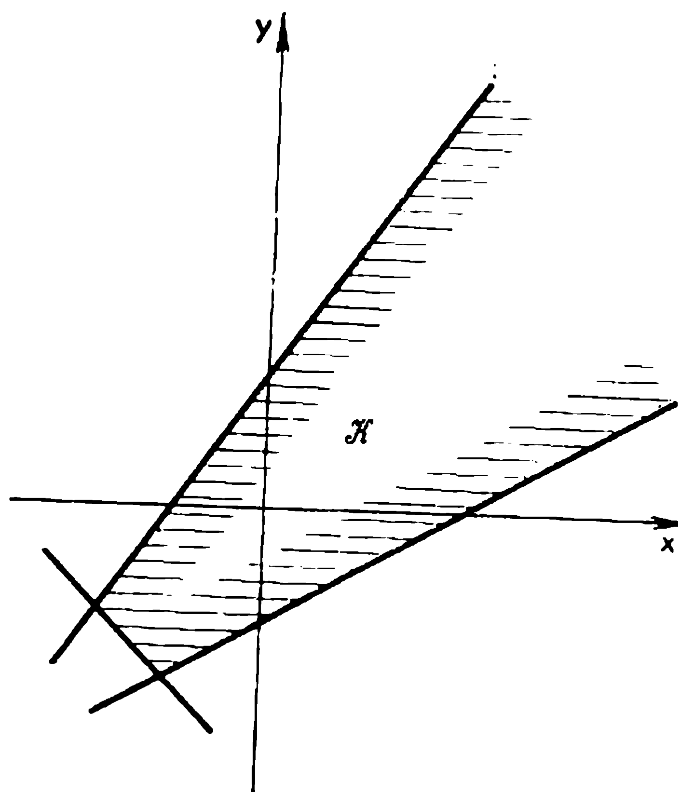


Fig. 17

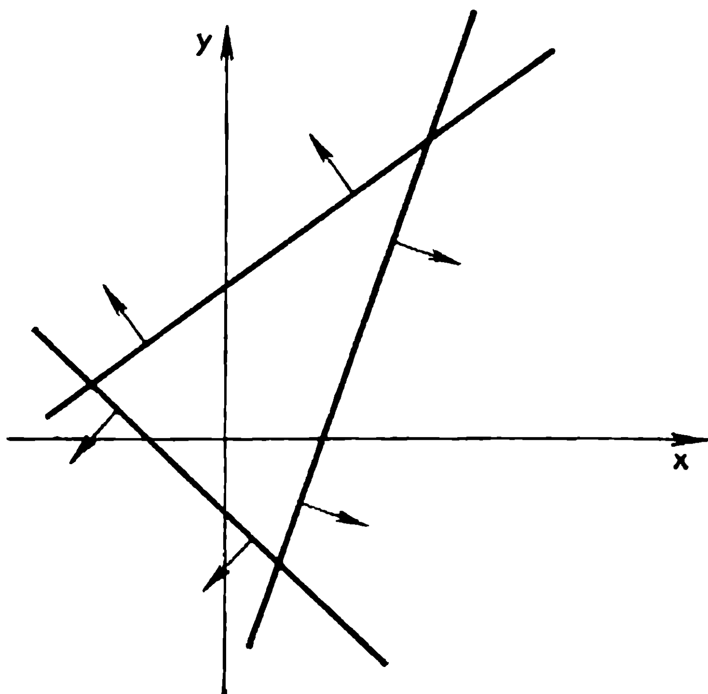
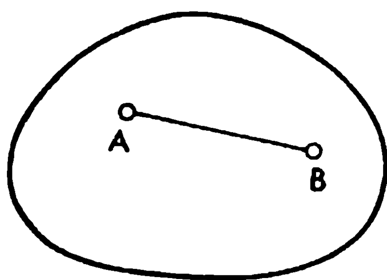


Fig. 18



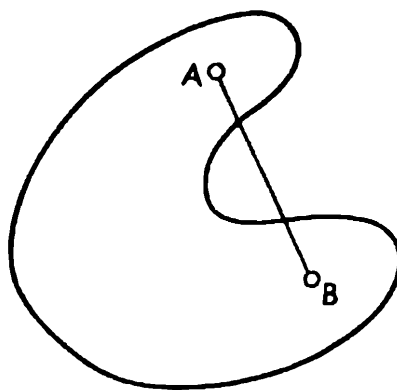


Fig. 19

deux points A et B quelconques il contient également tout le segment AB. La figure 19 montre la différence qui existe entre les ensembles convexe et non convexe. La convexité du domaine de solutions $\mathscr{K}$ découle du procédé même de formation de ce domaine; ce dernier s'obtient donc par intersection de plusieurs demi-plans dont chacun est un ensemble convexe.

D'ailleurs, pour éviter toute ambiguïté relative à la convexité du domaine $\mathscr{K}$, nous allons démontrer le lemme suivant.

● LEMME. *L'intersection d'un nombre quelconque d'ensembles convexes est aussi un ensemble convexe.*

● DÉMONSTRATION. Soient $\mathscr{K}_1$ et $\mathscr{K}_2$ deux ensembles convexes et $\mathscr{K}$ leur intersection. Considérons deux points A et B quelconques appartenant à $\mathscr{K}$ (fig. 20). Vu que $A \in \mathscr{K}_1$, $B \in \mathscr{K}_1$ et l'ensemble $\mathscr{K}_1$ est convexe, le segment AB appartient lui aussi à $\mathscr{K}_1$. De façon analogue le segment AB appartient à $\mathscr{K}_2$. Ainsi, le segment AB appartient simultanément à $\mathscr{K}_1$ et $\mathscr{K}_2$, par conséquent, il appartient à leur intersection $\mathscr{K}$. Nous avons ainsi démontré que $\mathscr{K}$ est un ensemble convexe. Le raisonnement analogue montre que l'intersection d'un nombre arbitraire (pas obligatoirement de deux ensembles) d'ensembles convexes est un ensemble convexe.

Ainsi, *le lieu géométrique des points dont les coordonnées vérifient toutes les inégalités* (1), *ou le domaine de solutions du système* (1), *est un domaine convexe polygonal $\mathscr{K}$. Il s'obtient par suite de l'intersection de tous les demi-plans correspondant aux inégalités du système considéré.*

Passons maintenant à un système à trois inconnues. Soit un système

$$\left.\begin{array}{r} a_1x + b_1y + c_1z + d_1 \geqslant 0, \\ a_2x + b_2y + c_2x + d_2 \geqslant 0, \\ \dots\dots\dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz + d_m \geqslant 0. \end{array}\right\} \qquad (2)$$

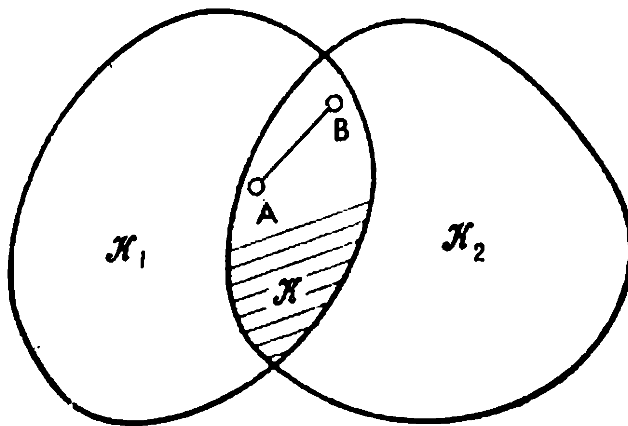


Fig. 20

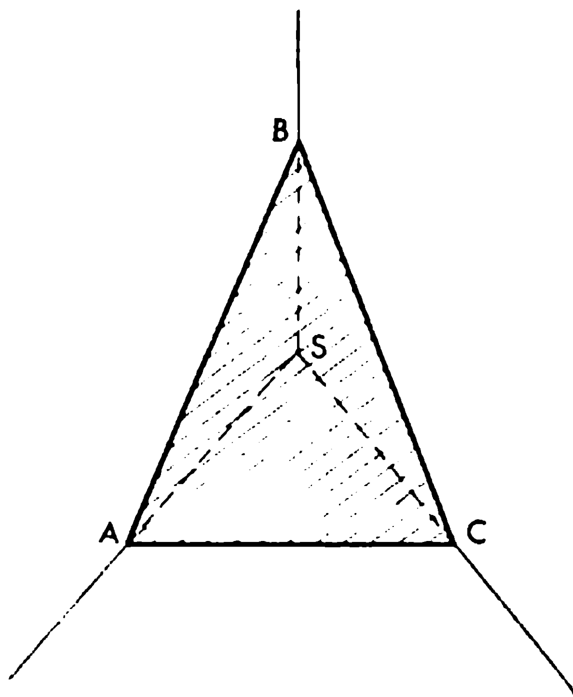


Fig. 21

Chacune de ces inégalités définit, comme nous le connaissons du § 1, un certain demi-espace. Pour cette raison, le domaine défini par ce système représente l'intersection (la partie commune) de m demi-espaces. Mais l'intersection d'un nombre fini de demi-espaces est un domaine polyèdre convexe $\mathscr{K}$. La figure 21 montre un tel domaine pour $m = 4$. Ici le domaine $\mathscr{K}$ est un tétraèdre ordinaire (ou, plus précisément, $\mathscr{K}$ comprend tous les points situés à l'intérieur ou sur la frontière de celui-ci). On comprend facilement que tout polyèdre convexe peut s'obtenir par l'intersection d'un nombre fini de demi-espaces *). Certes, le cas est possible où le domaine $\mathscr{K}$ n'est pas limité (il s'étend à l'infini), comme c'est le cas représenté sur la figure 22. Il peut, enfin, s'avérer qu'il n'existe pas en général de points vérifiant toutes les inégalités considérées (le système (2) est contradictoire) ; dans ce cas le domaine $\mathscr{K}$ est vide. C'est le cas représenté sur la figure 23.

Arrêtons-nous tout particulièrement sur le cas où parmi les inégalités (2) il y en a deux de la forme :

$$ax + by + cz + d \geqslant 0,$$
$$-ax - by - cz - d \geqslant 0,$$

qui peuvent être remplacées par une seule *équation*

$$ax + by + cz + d = 0.$$

Celle-ci définit dans l'espace un certain plan π. Les autres inégalités (2) prélèveront à partir du plan π un certain

---

*) Il convient de faire ici une remarque analogue à celle faite à la page 316. Le fait est que dans le programme scolaire de géométrie on entend par « polyèdre » une *surface* fermée composée de faces planes. Dans ce qui suit nous allons comprendre ce terme dans un sens plus large, en entendant par « polyèdre » non pas la surface elle-même, mais l'*ensemble de tous les points de l'espace qu'elle renferme* (bien entendu, cet ensemble contient également la surface elle aussi, mais en tant que sa partie).

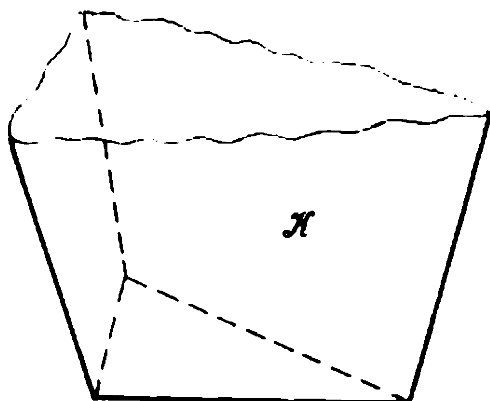
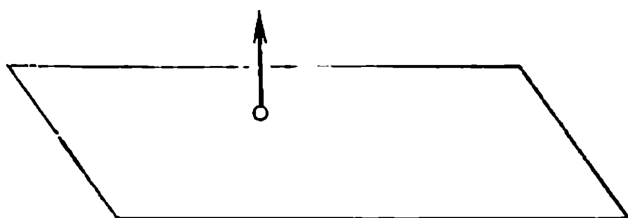


Fig. 22



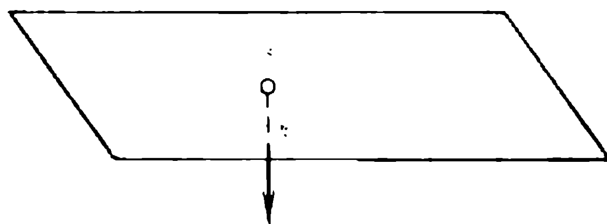


Fig. 23

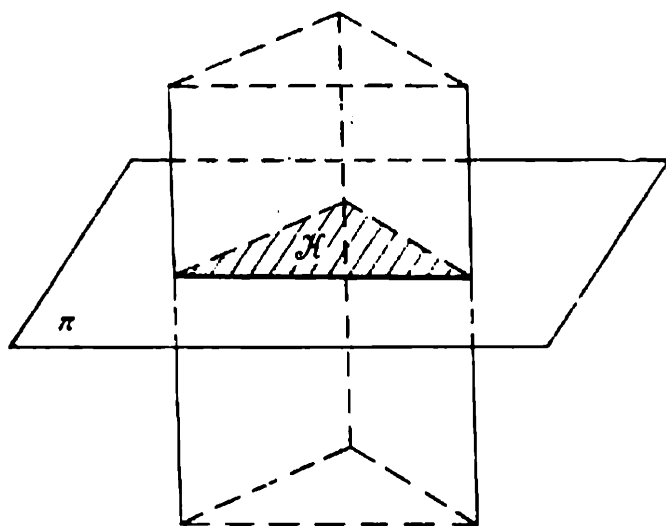


Fig. 24

domaine polygonal convexe qui sera justement pour le système (2) le domaine de solutions. Nous voyons que *le domaine polygonal convexe dans un plan peut constituer le cas particulier d'un domaine polyèdre convexe dans un espace.* Le domaine $\mathscr{K}$ représenté sur la figure 24 est un triangle formé par l'intersection de cinq demi-espaces dont deux sont limités par le plan horizontal π et les trois autres demi-espaces forment dans leur intersection un prisme trièdre « vertical ».

Par analogie avec le cas du système à deux inconnues, le domaine $\mathscr{K}$ s'appelle *domaine de solutions du système* (2). Il est à souligner une fois de plus que le domaine $\mathscr{K}$, qui est l'intersection d'un nombre quelconque de demi-espaces, est nécessairement convexe.

Ainsi, *le système* (2) *définit dans l'espace un domaine polyèdre convexe* $\mathscr{K}$. *Ce domaine s'obtient de l'intersection de tous les demi-espaces correspondant aux inégalités du système considéré.*

Si le domaine $\mathscr{K}$ est limité on l'appelle *polyèdre de solutions du système* (1).

§ 3. ENVELOPPE CONVEXE D'UN SYSTÈME DE POINTS

Imaginons-nous maintenant qu'en des points $A_1, A_2, \ldots$ $\ldots, A_p$ d'un plan représentant une feuille infinie d'un contre-plaqué on a enfoncé des clous. Prenons un boucle en caoutchouc, détendons-le bien et embrassons par ce dernier tous les clous (ligne en pointillé sur la fig. 25). Laissons ensuite à ce boucle le temps de se resserrer jusqu'à ce que les clous le permettent. L'ensemble des points embrassés par le boucle après son serrage est marqué de hachures sur la figure 25. Cet ensemble représente évidemment un certain polygone convexe qui s'appelle *enveloppe convexe d'un système de points* $A_1, A_2, \ldots, A_p$.

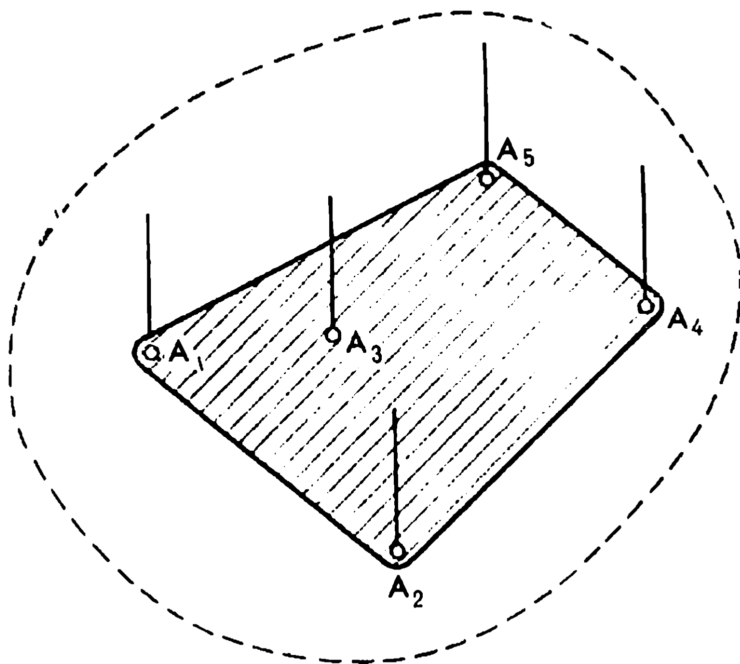


Fig. 25

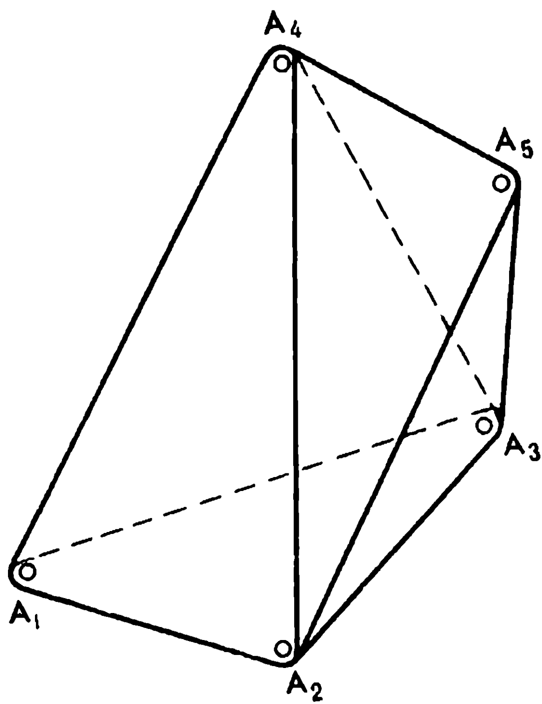


Fig. 26

Si les points $A_1, A_2, \ldots, A_p$ se trouvent non pas dans un plan mais dans un espace, il sera beaucoup plus difficile de réaliser une expérience analogue. Cependant, imaginons-nous tout de même qu'on a réussi à mettre tous les points $A_1, A_2, \ldots, A_p$ dans un sac en caoutchouc très élastique. Le sac élastique se serrera progressivement jusqu'à ce que certains de ces clous ne l'empêchent. Enfin, un moment arrivera où le serrage cessera (fig. 26). Il est aisé de comprendre qu'en ce moment le sac prendra la forme d'un polyèdre convexe avec les sommets en des points $A_1, A_2, \ldots, A_p$. Le domaine de l'espace embrassé par ce polyèdre sera également dit enveloppe convexe d'un système de points $A_1, A_2, \ldots, A_p$.

La définition de l'enveloppe convexe énoncée ci-dessus étant assez concrète n'est en revanche assez parfaite du point de vue mathématique. Donnons maintenant cette définition sous la forme plus stricte.

Soit $A_1, A_2, \ldots, A_p$ une collection arbitraire de points (dans un plan ou dans un espace). Examinons les points de la forme

$$s_1A_1 + s_2A_2 + \ldots + s_pA_p, \tag{1}$$

où $s_1, s_2, \ldots, s_p$ sont des nombres non négatifs quelconques dont la somme est égale à l'unité:

$$s_1, s_2, \ldots, s_p \geqslant 0 \text{ et } s_1 + s_2 + \ldots + s_p = 1. \tag{2}$$

● DÉFINITION. *L'ensemble des points de la forme* (1) *observant la condition* (2) *s'appelle enveloppe convexe du système de points* $A_1, A_2, \ldots, A_p$ *et est noté*

$$\langle A_1, A_2, \ldots, A_p \rangle.$$

Pour se convaincre de ce que cette définition ne diffère pas de celle donnée ci-dessus, nous examinerons d'abord les deux cas suivants: $p = 2$ et $p = 3$. Si $p = 2$, nous avons deux points A_1 et A_2. L'ensemble $\langle A_1, A_2 \rangle$ est, selon la proposition 1) du § 1, le segment A_1A_2.

Si $p = 3$, nous avons trois points A_1, A_2 et A_3. Montrons que l'ensemble $\langle A_1, A_2, A_3 \rangle$ se compose de tous les points situés à l'intérieur et sur les côtés du triangle $A_1A_2A_3$.
Pour cela démontrons le lemme suivant.

● LEMME. *L'ensemble $\langle A_1, \ldots, A_{p-1}, A_p \rangle$ se compose des segments qui relient le point A_p avec les points de l'ensemble $\langle A_1, \ldots, A_{p-1} \rangle$.*

● DÉMONSTRATION. Pour simplifier les notations, nous désignerons par $\mathscr{M}_{p-1}$ l'ensemble $\langle A_1, \ldots, A_{p-1} \rangle$ et par $\mathscr{M}_p$ l'ensemble $\langle A_1, \ldots, A_{p-1}, A_p \rangle$.
Examinons un point quelconque $A \in \mathscr{M}_p$. Il est de la forme

$$A = s_1A_1 + \ldots + s_{p-1}A_{p-1} + s_pA_p,$$

où

$$s_1, \ldots, s_p \geqslant 0, \ s_1 + \ldots + s_p = 1.$$

Si $s_p = 0$, on a $A \in \mathscr{M}_{p-1}$; de cette façon l'ensemble $\mathscr{M}_{p-1}$ fait partie de $\mathscr{M}_p$. Si $s_p = 1$, alors $A = A_p$; par suite, le point A_p appartient à $\mathscr{M}_p$. Ainsi, $\mathscr{M}_p$ contient $\mathscr{M}_{p-1}$ de même que le point A_p. Montrons maintenant que *tout segment $A'A_p$, où $A' \in \mathscr{M}_{p-1}$, appartient entièrement à $\mathscr{M}_p$.*
Si A est un point d'un tel segment, nous avons

$$A = tA' + sA_p \quad (t, s \geqslant 0, \ t + s = 1).$$

D'autre part, d'après la définition du point A', on a

$$A' = t_1A_1 + \ldots + t_{p-1}A_{p-1}$$

$$(t_1, \ldots, t_{p-1} \geqslant 0, \ t_1 + \ldots + t_{p-1} = 1);$$

par conséquent,

$$A = tt_1A_1 + \ldots + tt_{p-1}A_{p-1} + sA_p.$$

En posant $tt_1 = s_1, \ldots, tt_{p-1} = s_{p-1}, s = s_p$, on obtient (1), (2). Ainsi, nous avons démontré que $A \in \mathscr{M}_p$. Nous pou-

vons donc en conclure que n'importe lequel des segments indiqués ci-dessus appartient entièrement à $\mathscr{M}_p$.

Il nous reste à vérifier que l'ensemble $\mathscr{M}_p$ ne contient que les segments en question, c'est-à-dire que *tout point A de $\mathscr{M}_p$ appartient à l'un des segments considérés.*

Soit $A \in \mathscr{M}_p$. Nous avons alors (1), (2). On peut de plus considérer que $s_p \neq 1$, autrement $A = A_p$ et donc il n'y a rien à démontrer. Mais si $s_p \neq 1$, on a $s_1 + \ldots + s_{p-1} = 1 - s_p > 0$ et on peut écrire

$$A = (s_1 + \ldots + s_{p-1}) \left[\frac{s_1}{s_1 + \ldots + s_{p-1}} A_1 + \ldots \right.$$

$$\left. \ldots + \frac{s_{p-1}}{s_1 + \ldots + s_{p-1}} A_{p-1} \right] + s_p A_p.$$

L'expression entre crochets définit un certain point A' appartenant à $\mathscr{M}_{p-1}$, étant donné que les coefficients de $A_1, \ldots$ $\ldots, A_{p-1}$ sont non négatifs et leur somme vaut 1. Ainsi,

$$A = (s_1 + \ldots + s_{p-1}) A' + s_p A_p.$$

Du moment que les coefficients de A' et A_p sont également non négatifs et leur somme vaut 1, le point A appartient au segment $A'A_p$. Le lemme est donc démontré.

Il est aisé de comprendre que l'image concrète de l'enveloppe convexe donnée au début de ce paragraphe et la définition stricte ultérieure sont équivalentes. En effet, quelle que soit la définition que nous chosissons, le passage de l'enveloppe convexe du système $A_1, \ldots, A_{p-1}$ à l'enveloppe convexe du système $A_1, \ldots, A_{p-1}, A_p$ s'effectue suivant la même règle: il faut relier le point A_p par des segments avec tous les points de l'enveloppe convexe du système $A_1, \ldots, A_{p-1}$ (dans le cas de la définition concrète de l'enveloppe convexe cette règle est évidente et dans le cas de la définition stricte cette règle constitue le sens du lemme démontré). Si l'on tient compte maintenant que selon les deux définitions pour $p = 2$ nous obtenons le même

ensemble, le segment A_1A_2, l'équivalence de ces deux définitions deviendra évidente.

Le terme « enveloppe convexe » n'est d'ailleurs pas tout à fait justifié puisque nous n'avons pas encore démontré que l'*ensemble* $\langle A_1, A_2, \ldots, A_p \rangle$ *est toujours convexe.* Nous allons le montrer.

Soient A et B deux points arbitraires de cet ensemble:

$$A = s_1A_1 + s_2A_2 + \ldots + s_pA_p,$$
$$B = t_1A_1 + t_2A_2 + \ldots + t_pA_p,$$

où

$$s_1, \ldots, s_p, t_1, \ldots, t_p \geqslant 0,$$
$$s_1 + \ldots + s_p = t_1 + \ldots + t_p = 1. \qquad (3)$$

Tout point C du segment AB est de la forme

$$C = sA + tB,$$

où

$$s, t \geqslant 0, \ s + t = 1, \qquad (4)$$

d'où il s'ensuit

$$C = s(s_1A_1 + \ldots + s_pA_p) + t(t_1A_1 + \ldots + t_pA_p) =$$
$$= (ss_1 + tt_1) A_1 + \ldots + (ss_p + tt_p) A_p.$$

Les coefficients de $A_1, \ldots, A_p$ sont non négatifs et leur somme vaut 1 (ceci découle de (3), (4)). Ceci montre que le point C appartient à l'ensemble $\langle A_1, A_2, \ldots, A_p \rangle$, c'est-à-dire cet ensemble est convexe.

Il est également aisé de voir que l'ensemble $\langle A_1, A_2, \ldots, A_p \rangle$ est *le plus petit de tous les ensembles convexes qui contiennent les points initiaux* $A_1, A_2, \ldots, A_p$, c'est-à-dire il est contenu dans n'importe lequel de ces ensembles. Cette assertion découle directement du lemme démontré plus haut et de la définition de l'ensemble convexe.

Le fait ci-dessus permet d'expliquer le terme « enveloppe convexe » et encore que l'ensemble $\langle A_1, A_2, \ldots, A_p \rangle$ peut

s'obtenir à l'aide du procédé décrit au début de ce paragraphe. En effet, l'ensemble enveloppé par le boucle élastique (ou la pellicule) après son resserrement maximal autour du système de points $A_1, A_2, \ldots, A_p$ est justement le plus petit ensemble convexe renfermant les points considérés.

§ 4. CÔNE POLYÈDRE CONVEXE

Nous commençons par une définition.

On appelle cône polyèdre convexe l'intersection d'un nombre fini de demi-espaces dont les plans limites passent par un point commun appelé sommet du cône.

Définissons d'abord la corrélation qui existe entre la notion de cône polyèdre convexe et les systèmes d'inégalité linéaires. Commençons par le cas particulier, notamment, celui où le sommet du cône se confond avec l'origine des coordonnées. Cela signifie que tous les plans limites passent par l'origine des coordonnées. Mais l'équation d'un plan passant par l'origine est de la forme

$$ax + by + cz = 0$$

(le terme libre de l'équation doit être nul, sinon (0, 0, 0) ne sera pas une solution). Ainsi, *un cône polyèdre convexe dont le sommet coïncide avec l'origine des coordonnées est le domaine de solutions d'un système d'inégalités homogènes*:

$$\left.\begin{array}{l} a_1x + b_1y + c_1z \geqslant 0, \\ a_2x + b_2y + c_2z \geqslant 0, \\ \dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz \geqslant 0. \end{array}\right\}$$

Il est évident que la réciproque est vraie: le domaine de solutions d'un système homogène d'inégalités représente toujours un certain cône polyèdre convexe ayant l'origine des coordonnées pour son sommet.

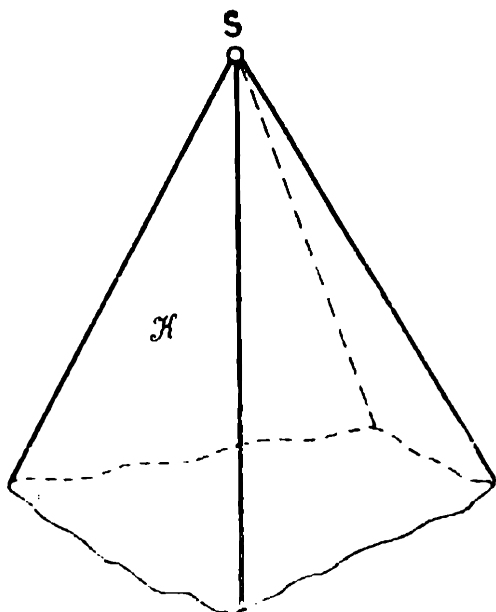


Fig. 27

A titre d'exemple d'un cône polyèdre convexe on peut citer un domaine convexe dans un espace, ayant pour sa limite un certain angle polyèdre de sommet S, une sorte de pyramide convexe illimitée sans base et s'étendant d'une façon illimitée à partir de son sommet (comme cette pyramide tétraédrique représentée sur la fig. 27). On peut également citer d'autres exemples bien que moins intéressants :

1. Un demi-espace (fig. 28,*a*). Un tel « cône » peut avoir pour son sommet tout point $S \in \pi$, π étant le plan limite du demi-espace considéré.

2. L'intersection de deux demi-espaces dont les plans limites se coupent suivant une certaine droite l (fig. 28,*b*). Pour sommet on peut prendre tout point $S \in l$.

3. Un plan. Il est clair qu'on peut considérer tout plan π dans l'espace comme l'intersection de deux demi-espaces situés de deux côtés de π (fig. 28,*c*). Dans ce cas tout point $S \in \pi$ peut être considéré comme son sommet.

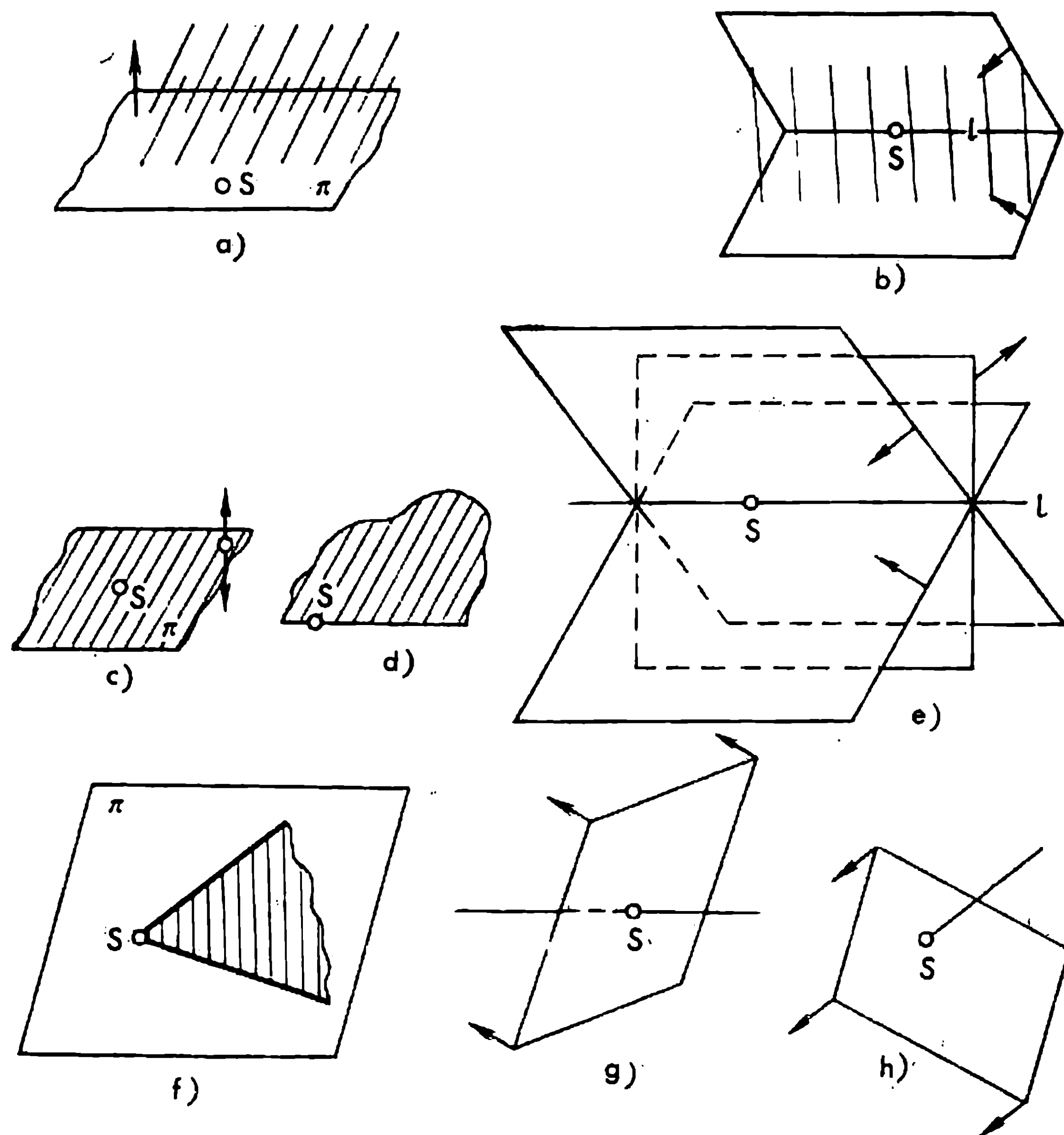


Fig. 28

4. Un demi-espace (fig. 28,d). N'importe quel point de la droite limite joue le rôle de son sommet S.

5. Une droite. Chaque droite l dans l'espace peut s'obtenir par l'intersection de trois demi-espaces dont les plans limites passent par l (fig. 28,e). Le sommet S est un point arbitraire de la droite l.

6. Un angle (inférieur à 180°) dans un plan arbitraire π (fig. 28,f). Un tel angle s'obtient si l'on coupe le plan π par deux demi-espaces (de quelle façon?).

7. Un rayon (fig. 28,g). Il peut être considéré comme l'intersection d'une droite avec un demi-espace. Le sommet S coïncide avec l'origine du rayon.

8. Un point. Un tel « cône » s'obtient si l'on prend la partie commune d'un rayon et d'un demi-espace correspondant (fig. 28,h).

Certes, les exemples cités 1-8 diffèrent du sens habituel du terme « cône », mais nous devons le tolérer si nous voulons conserver la définition générale du cône polyèdre convexe donnée au début de ce paragraphe.

Essayons maintenant de montrer que *tous les cônes polyèdres convexes dans un espace se réduisent aux ensembles mentionnés plus haut.*

Soit p le nombre de demi-espaces dont l'intersection forme un cône considéré $\mathscr{K}$. Si $p = 1$, notre assertion est juste, car $\mathscr{K}$ représente alors un demi-espace. Nous laissons au lecteur le soin de montrer, à l'aide de raisonnements fort simples, que si cette assertion est juste pour un cône obtenu par l'intersection de p demi-espaces, elle le sera aussi pour celui formé de $p + 1$ demi-espaces. D'où, selon le principe de récurrence, il s'ensuit que notre assertion sera valide pour tout p.

Les cônes polyèdres convexes jouissent de propriétés très intéressantes. Malheureusement, nous ne pouvons pas

nous attarder ici sur les détails, et quand même nous tâcherons d'exposer quelques faits intéressants dans ce paragraphe et dans le § 9.

Introduisons encore une définition, ou plutôt une notation.

Soit $B_1, B_2, \ldots, B_q$ une collection arbitraire d'un nombre fini de points (dans un espace). *Nous désignerons par* $(B_1, B_2, \ldots, B_q)$ *l'ensemble des points de la forme*

$$t_1B_1 + t_2B_2 + \ldots + t_qB_q,$$

où $t_1, t_2, \ldots, t_q$ *sont des nombres quelconques non négatifs.*

Qu'est-ce que représente du point de vue géométrique l'ensemble $(B_1, B_2, \ldots, B_q)$? De la définition il résulte que celui-ci est une somme des ensembles $(B_1), (B_2), \ldots, (B_q)$; il nous faut donc commencer par élucider le sens de l'ensemble (B), c'est-à-dire l'ensemble des points de la forme tB, où t est un nombre arbitraire non négatif et B un point fixe. La réponse à cette dernière question est évidente: si B est l'origine des coordonnées, l'ensemble (B) se confond également avec celui-ci; sinon (B) représente un rayon issu de l'origine des coordonnées et passant par le point B. Il est à remarquer que la somme d'un ensemble quelconque avec l'origine des coordonnées est de nouveau ce même ensemble, d'où il s'ensuit qu'en examinant les ensembles de la forme $(B_1, B_2, \ldots, B_q)$ *nous pouvons donc considérer tous les points* $B_1, B_2, \ldots, B_q$ *comme ne coïncidant pas avec l'origine des coordonnées. L'ensemble* $(B_1, B_2, \ldots, B_q)$ *représentera dans ce cas la somme des rayons* $(B_1), (B_2), \ldots, (B_q)$.

Cette dernière remarque rend presque évident le lemme suivant.

● LEMME *L'ensemble* $(B_1, \ldots, B_{q-1}, B_q)$ *est la réunion des segments reliant chaque point de l'ensemble* $(B_1, \ldots, B_{q-1})$ *avec chaque point du rayon* (B_q).

La démonstration rigoureuse de ce lemme étant parfaitement analogue à celle du lemme du § 3, nous laissons donc au lecteur le soin de le faire personnellement.

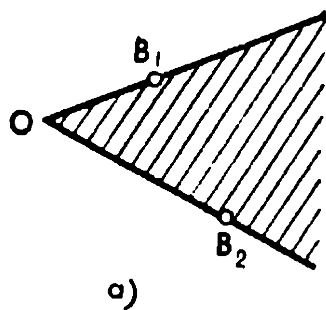
En partant de ce lemme, il est facile de déduire que (B_1, B_2) est un angle, une droite ou un rayon (fig. 29,*a*, *b*, *c*,). On établit ensuite que (B_1, B_2, B_3) est l'un des ensembles suivants: une pyramide trièdre illimitée, un plan, un demi-plan, un angle, une droite, un rayon. Il devient maintenant clair qu'entre les ensembles $(B_1, B_2, \ldots, B_q)$ et les cônes polyèdres convexes doit exister une relation. Et cette relation existe effectivement. Pour plus d'évidence nous énoncerons les propositions correspondantes sous forme de deux théorèmes suivants.

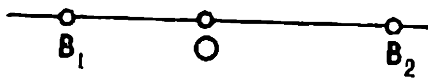
● THÉORÈME 1. *L'ensemble* $(B_1, B_2, \ldots, B_q)$ *soit coïncide avec tout l'espace soit représente un cône polyèdre convexe ayant pour sommet l'origine des coordonnées.*

L'exemple suivant montre que l'ensemble $(B_1, B_2, \ldots$ $\ldots, B_q)$ peut effectivement coïncider avec tout l'espace. Examinons quatre points B_1, B_2, B_3, B_4 disposés de façon que les rayons (B_1), (B_2), (B_3), (B_4) constituent deux à deux des angles obtus (fig. 30). Chacun des ensembles (B_1, B_2, B_3), (B_1, B_2, B_4), (B_1, B_3, B_4), (B_2, B_3, B_4) représente une pyramide trièdre infinie de sommet à l'origine des coordonnées. L'ensemble (B_1, B_2, B_3, B_4) renferme, évidemment, chacune de ces pyramides. Mais la réunion des pyramides en question coïncide avec tout l'espace!

● THÉORÈME 2. *Tout cône polyèdre convexe de sommet à l'origine des coordonnées représente un ensemble de la forme* $(B_1, B_2, \ldots, B_q)$.

Nous n'effectuerons la *démonstration du théorème 1* que dans les grandes lignes. Faisons pour cela recours à la méthode de récurrence. L'assertion du théorème pour $q = 1$ est





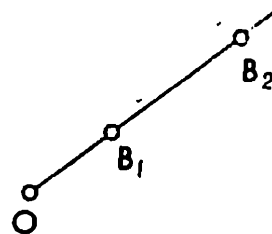


Fig. 29

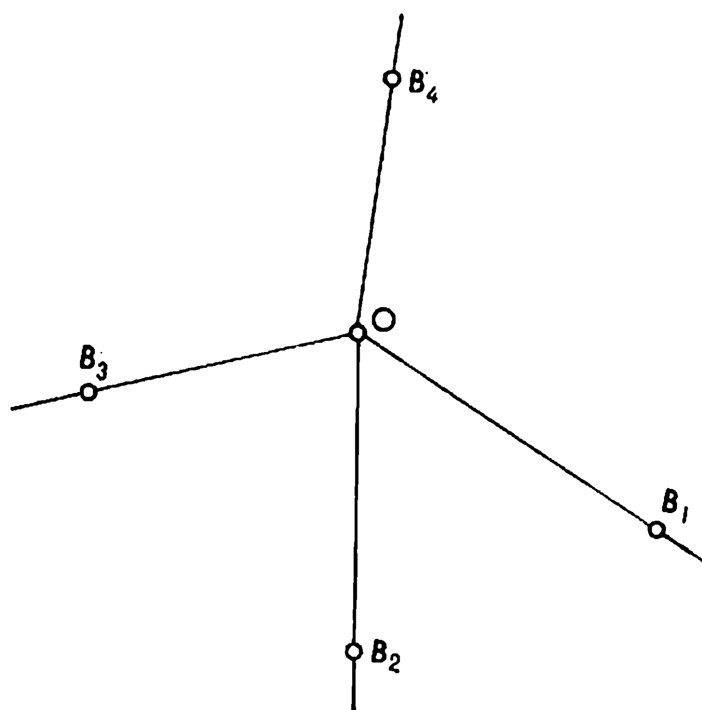


Fig. 30

évidente. Supposons maintenant que ce théorème soit vrai pour un ensemble de la forme $(B_1, \ldots, B_q)$. Nous appuyant sur ce fait, essayons de démontrer la validité de ce théorème pour les ensembles $(B_1, \ldots, B_q, B_{q+1})$.

Par récurrence $(B_1, \ldots, B_q)$ est ou bien tout l'espace, ou bien un cône polyèdre convexe contenu dans celui-ci. Dans le premier cas il n'y a rien à démontrer, car autrement $(B_1, \ldots, B_q, B_{q+1})$ sera alors lui aussi tout l'espace. Examinons donc le deuxième cas: $(B_1, \ldots, B_q)$ est un cône polyèdre convexe $\mathcal{K}$. D'après le lemme démontré ci-dessus, l'ensemble $(B_1, \ldots, B_q, B_{q+1})$ est une réunion des segments reliant chaque point de l'ensemble $\mathcal{K}$ avec chaque point du rayon (B_{q+1}). Mais, comme nous l'avons montré plus haut, tout cône polyèdre convexe $\mathcal{K}$ est ou bien une pyramide convexe infinie, ou bien l'un des ensembles 1-8. Après avoir examiné pour chacun de ces cas la réunion considérée des segments, on se convainc facilement (nous laissons au lecteur le soin de le faire personnellement !) que cette réunion ou bien coïncide avec tout l'espace, ou bien devient de nouveau un cône polyèdre convexe. Ainsi, le théorème est vrai pour les ensembles de la forme (B_1) ainsi que pour ceux de la forme $(B_1, \ldots, B_q, B_{q+1})$ s'il l'est aussi pour $(B_1, \ldots, B_q)$. D'où l'on déduit que le théorème est vrai pour tout q.

Démonstration du théorème 2. Soit $\mathcal{K}$ un cône polyèdre convexe de sommet à l'origine des coordonnées O. Comme nous l'avons déjà dit, $\mathcal{K}$ représente soit une pyramide convexe illimitée, soit un des ensembles 1-8.

Soit $\mathcal{K}$ une pyramide. Choisissons sur chacune de ses arêtes un point quelconque, nous obtiendrons donc un système de points $B_1, B_2, \ldots, B_q$. Nous affirmons que *l'ensemble $(B_1, B_2, \ldots, B_q)$ est justement $\mathcal{K}$.*

Pour le démontrer, examinons un plan π quelconque coupant toutes les arêtes de la pyramide $\mathcal{K}$. On obtient les points $B'_1, B'_2, \ldots, B'_q$ (fig. 31). Il est évident que

$$B'_1 = k_1 B_1. \ B'_2 = k_2 B_2, \ldots, B'_q = k_q B_q, \qquad (1)$$

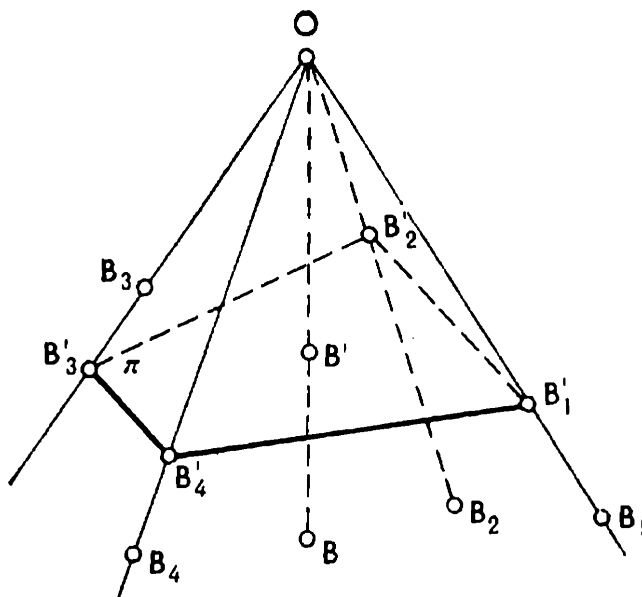


Fig. 31

où $k_1, k_2, \ldots, k_q$ sont des nombres non négatifs quelconques.

Soit maintenant un point quelconque B de la pyramide considérée différant du sommet O. Le rayon OB se coupe avec le plan π en un point B'.

Le point B' appartient évidemment à l'enveloppe convexe du système $B'_1, B'_2, \ldots, B'_q$, d'où

$$B' = s_1 B'_1 + s_2 B'_2 + \ldots + s_q B'_q,$$

où $s_1, s_2, \ldots, s_q$ sont des nombres non négatifs quelconques dont la somme vaut 1. Si l'on tient maintenant compte de (1), on obtient

$$B' = s_1 k_1 B_1 + s_2 k_2 B_2 + \ldots + s_q k_q B_q,$$

et si l'on prend en considération encore que $B' = kB$ ($k > 0$), on obtiendra définitivement

$$B = t_1 B_1 + t_2 B_2 + \ldots + t_q B_q,$$

où $t_i = \frac{s_i k_i}{k}$ ($i = 1, 2, \ldots, q$). Ainsi, nous avons montré que tout point B de la pyramide $\mathscr{K}$ appartient à l'ensemble

$(B_1, B_2, \ldots, B_q)$. La réciproque (c'est-à-dire tout point de l'ensemble $(B_1, B_2, \ldots, B_q)$ appartient à $\mathcal{K}$) est évidente. $\mathcal{K}$ coïncide donc avec $(B_1, B_2, \ldots, B_q)$.

Le cas où $\mathcal{K}$ est un des ensembles particuliers 1-8 étant fort élémentaire, nous laissons au lecteur le soin de le démontrer personnellement.

§ 5. DOMAINE DE SOLUTIONS D'UN SYSTÈME D'INÉGALITÉS À DEUX INCONNUES

Dans ce paragraphe, nous nous proposons comme but de donner une description efficace de toutes les solutions d'un système d'inégalités linéaires, plus précisément d'un système à deux inconnues x et y. Bien que le nombre d'inconnues soit restreint (il n'y en a que deux), nous tâcherons de faire l'analyse de ces systèmes à partir des positions générales pour que les résultats obtenus puissent être généralisés aux systèmes à un nombre d'inconnues plus grand.

La résolution de tout système d'inégalités linéaires se ramène en fin de compte à la résolution d'*une série des systèmes d'équations linéaires.* Nous considérerons la résolution d'un système d'équations linéaires comme une opération élémentaire sans nous embarrasser du fait qu'il nous faudra effectuer cette opération plusieurs fois.

A. *Lemmes nécessaires.* Soit

$$\left.\begin{aligned} a_1x + b_1y + c_1 &\geqslant 0, \\ a_2x + b_2y + c_2 &\geqslant 0, \\ \dots\dots&\dots\dots \\ a_mx + b_my + c_m &\geqslant 0 \end{aligned}\right\} \qquad (1)$$

un système d'inégalités linéaires. Il s'avère utile de considérer aussi le système correspondant d'*inégalités homo-*

gènes:

$$\left.\begin{aligned} a_1x + b_1y &\geqslant 0, \\ a_2x + b_2y &\geqslant 0, \\ \dots\dots& \\ a_mx + b_my &\geqslant 0, \end{aligned}\right\} \qquad (2)$$

ainsi que le système correspondant d'*équations homogènes*:

$$\left.\begin{aligned} a_1x + b_1y &= 0, \\ a_2x + b_2y &= 0, \\ \dots\dots& \\ a_mx + b_my &= 0. \end{aligned}\right\} \qquad (3)$$

Désignons par $\mathscr{K}$ le domaine de solutions du système (1) dans le plan des coordonnées xOy, par $\mathscr{K}_0$ celui du système (2) et par $\mathscr{L}$ du système (3). Il est clair que $\mathscr{L} \subset \mathscr{K}_0$ où le symbole $\subset$ signifie « inclus dans » *).

● LEMME. 1. *L'inclusion suivante*

$$\mathscr{K} + \mathscr{K}_0 \subset \mathscr{K}$$

a lieu, c'est-à-dire la somme de toute solution d'un système d'inégalités donné avec toute solution d'un système d'inégalités homogène correspondant est de nouveau une solution du système donné.

● DÉMONSTRATION. Soient A un point arbitraire de $\mathscr{K}$ et B un point quelconque de $\mathscr{K}_0$. Alors les inégalités suivantes

$$\begin{aligned} a_1x_A + b_1y_A + c_1 &\geqslant 0, \\ a_2x_A + b_2y_A + c_2 &\geqslant 0, \\ \dots\dots\dots& \\ a_mx_A + b_my_A + c_m &\geqslant 0 \end{aligned} \qquad \text{et} \qquad \begin{aligned} a_1x_B + b_1y_B &\geqslant 0, \\ a_2x_B + b_2y_B &\geqslant 0, \\ \dots\dots\dots& \\ a_mx_B + b_my_B &\geqslant 0 \end{aligned}$$

*) Il ne faut pas confondre le symbole $\subset$ avec celui $\in$ introduit plus haut. Ce dernier ne s'emploie que là où il s'agit de *l'appartenance d'un point quelconque à un ensemble quelconque*. Si l'on veut noter qu'un *ensemble quelconque est contenu dans un autre*, on emploie le symbole $\subset$.

sont vraies. En additionnant chaque inégalité de gauche avec l'inégalité correspondante de droite, on obtient

$$\begin{aligned} a_1 (x_A + x_B) + b_1 (y_A + y_B) + c_1 &\geqslant 0, \\ a_2 (x_A + x_B) + b_2 (y_A + y_B) + c_2 &\geqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots &\cdots \\ a_m (x_A + x_B) + b_m (y_A + y_B) + c_m &\geqslant 0. \end{aligned}$$

Ces inégalités montrent que le couple de nombres $x_A + x_B$, $y_A + y_B$, coordonnées du point $A + B$, est une solution du système initial (1), c'est-à-dire que $A + B \in \mathscr{K}$. Le lemme est donc démontré.

● LEMME 2. 1) *Si un rayon quelconque issu du point A appartient entièrement à l'ensemble $\mathscr{K}$ et P est un point arbitraire de ce rayon,* $P - A \in \mathscr{K}_0$.

2) *Si une droite quelconque appartient entièrement à l'ensemble $\mathscr{K}$ et A, P sont deux points arbitraires de cette droite,* $P - A \in \mathscr{L}$.

● DÉMONSTRATION 1) Désignons par B le point $P - A$. Le rayon considéré se compose de points de la forme:

$$A + sB, \tag{4}$$

où s est un nombre arbitraire non négatif (fig.32). N'importe lequel des points est, par hypothèse, une solution du système (1), c'est-à-dire

$$\left.\begin{aligned} a_1 (x_A + sx_B) + b_1 (y_A + sy_B) + c_1 &\geqslant 0, \\ a_2 (x_A + sx_B) + b_2 (y_A + sy_B) + c_2 &\geqslant 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots &\cdots \\ a_m (x_A + sx_B) + b_m (y_A + sy_B) + c_m &\geqslant 0. \end{aligned}\right\} \tag{5}$$

Examinons, par exemple, la première de ces inégalités. Elle peut s'écrire sous la forme

$$(a_1x_A + b_1y_A + c_1) + s\,(a_1x_B + b_1y_B) \geqslant 0.$$

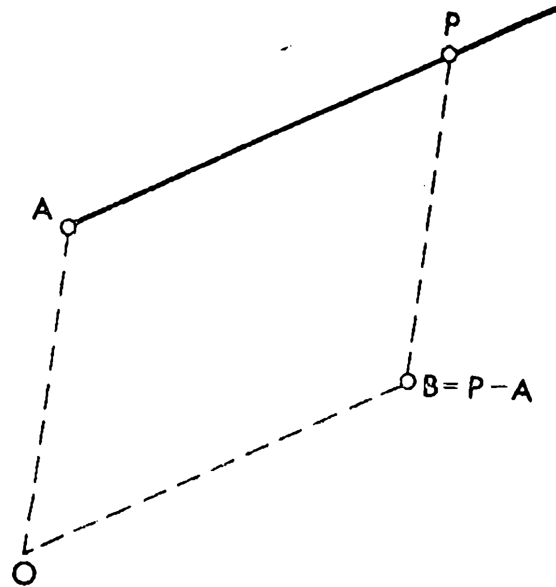


Fig. 32

Du moment que cette inégalité est vraie pour *tout* $s \geqslant 0$, le coefficient de s, comme on le voit aisément, doit être un nombre non négatif:

$$a_1 x_B + b_1 y_B \geqslant 0.$$

De façon analogue, les autres inégalités (5) nous donnent

$$a_2 x_B + b_2 y_B \geqslant 0,$$
$$\cdots\cdots\cdots$$
$$a_m x_B + b_m y_B \geqslant 0,$$

d'où il s'ensuit que le point B appartient à l'ensemble $\mathcal{K}_0$.

La démonstration de 2) s'effectue de façon absolument identique à celle de 1). La droite considérée se compose de points de la forme (4), où s est un nombre arbitraire. Pour cette raison, les inégalités (5) sont vraies pour tout s. D'où il s'ensuit que le coefficient sommaire de s dans chacune de ces inégalités doit être nul, c'est-à-dire

$$a_1 x_B + b_1 y_B = 0,$$
$$a_2 x_B + b_2 y_B = 0,$$
$$\cdots\cdots\cdots$$
$$a_m x_B + b_m y_B = 0.$$

Par conséquent, $B \in \mathcal{L}$. Le lemme est donc démontré.

Il est aisé de voir que les lemmes 1 et 2 sont vrais pour des systèmes à un nombre d'inconnues quelconque.

B. *Cas où le système d'inégalités* (1) *est normal.* Considérons de nouveau le système d'inégalités (1) et le système d'équations homogènes (3) qui lui correspond. Cette dernière possède, évidemment, une solution évidente $x = 0$, $y = 0$. Cette solution s'appelle *nulle*. Pour explorer le système (1), il s'avère important de savoir si le système (3) possède des solutions non nulles. Donnons à cet effet la définition suivante.

● DÉFINITION. *Un système d'inégalités linéaires est dit normal si le système correspondant des équations linéaires homogènes n'a qu'une solution nulle.*

En d'autres termes, un système d'inégalités est normal si l'ensemble $\mathscr{L}$, domaine de solutions du système d'équations homogène correspondant, ne contient qu'un seul point (origine des coordonnées).

Il va de soi que la notion de système normal a certainement le sens pour un nombre quelconque d'inconnues.

Il est aisé de voir qu'un *système d'inégalités compatible est normal si et seulement si le domaine de ses solutions* $\mathscr{K}$ *ne contient aucune droite.*

En effet, si un système est normal, c'est-à-dire l'ensemble $\mathscr{L}$ ne contient que l'origine des coordonnées, le domaine $\mathscr{K}$ ne contient aucune droite — ceci découle immédiatement de la deuxième assertion du lemme 2. Par contre, si un système n'est pas normal, l'ensemble $\mathscr{L}$ renferme alors au moins un point B ne coïncidant pas avec l'origine des coordonnées. Certes, tous les points de la forme kB, k étant un nombre arbitraire, appartiennent également à $\mathscr{L}$ *). Mais dans ce cas quel que soit un point $P \in \mathscr{K}$ (un tel point existe

*) Si les nombres x, y, z, coordonnées du point B, satisfont à un système d'équations homogène, les nombres kx, ky, kz, coordonnées du point kB, le font aussi.

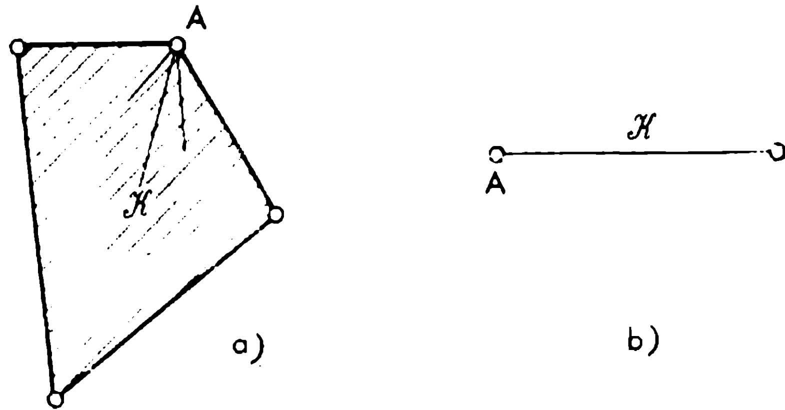


Fig. 33

toujours puisque le système en question est compatible et, par conséquent, le domaine $\mathcal{K}$ n'est pas vide), l'ensemble de tous les points de la forme $P + kB$ (k étant un nombre quelconque) appartient, selon le lemme 1, à $\mathcal{K}$. Cet ensemble, comme nous le savons, est une droite. Par conséquent, si le système considéré n'est pas normal, le domaine $\mathcal{K}$ contient une droite. Ceci démontre la proposition formulée ci-dessus.

Nous passons maintenant à l'étude du domaine de solutions du système normal (1) tout en supposant ce système compatible (le domaine $\mathcal{K}$ n'est pas vide) et normal.

Tout d'abord, le fait que le domaine $\mathcal{K}$ ne contient aucune droite conduit à ce qu'il doit nécessairement avoir des sommets. A la notion de sommet nous attribuons le sens suivant:

On appelle sommet d'un domaine $\mathcal{K}$ un point de ce domaine tel qui n'est pas intérieur par rapport à aucun segment entièrement contenu dans $\mathcal{K}$. En d'autres termes, le sommet est un point $A \in \mathcal{K}$ possédant la propriété suivante: tout segment appartenant à $\mathcal{K}$ et passant par le point A doit avoir en ce point son origine ou son extrémité (fig. 33,*a* et *b*, où le point A est l'un des sommets; fig. 33,*b* où le domaine $\mathcal{K}$ est un segment).

Expliquons de façon plus détaillée pourquoi l'ensemble convexe $\mathcal{K}$ qui nous intéresse possède des sommets. Si $\mathcal{K}$ est situé sur une droite, il représente alors soit un point

isolé, soit un segment, soit enfin un rayon, et l'existence des sommets est donc évidente. Si $\mathcal{K}$ n'est pas situé sur une droite, nous examinerons la frontière de cet ensemble. Celle-ci est constituée de segments et de rayons ($\mathcal{K}$ ne contient pas de droites entières). L'extrémité de n'importe lequel de ces segments et l'origine de tout rayon seront évidemment les sommets de $\mathcal{K}$.

La détermination des sommets du domaine $\mathcal{K}$ ne présente aucune difficulté. Remarquons tout d'abord qu'à la i-ème inégalité du système (1) correspond dans le plan des coordonnées xOy un demi-plan dont la droite limite l_i a pour équation

$$a_i x + b_i y + c_i = 0 \quad (i = 1, 2, \ldots, m).$$

Il est clair que le point A du domaine $\mathcal{K}$ est le sommet de ce dernier si et seulement s'il appartient à deux droites limites *différentes.*

Convenons d'appeler *régulier* tout sous-système de deux équations du système

$$\left.\begin{aligned} a_1 x + b_1 y + c_1 &= 0, \\ a_2 x + b_2 y + c_2 &= 0, \\ \ldots\ldots\ldots\ldots \\ a_m x + b_m y + c_m &= 0 \end{aligned}\right\} \qquad (6)$$

à condition que ce sous-système n'ait qu'une *seule* solution (x, y).

Il résulte de la caractéristique d'un sommet exposée ci-dessus le procédé suivant de détermination des sommets d'un domaine $\mathcal{K}$.

Pour déterminer tous les sommets du domaine $\mathcal{K}$ il faut trouver les solutions de tous les sous-systèmes réguliers du système (6) *pour en choisir ensuite celles qui vérifient le système initial (1).*

Etant donné que le nombre de sous-systèmes réguliers ne dépasse pas C_m^2, nombre de combinaisons de m pris 2 à 2,

le nombre de sommets du domaine $\mathscr{K}$ ne peut non plus être supérieur. *Ainsi, le nombre de sommets est fini.*

● REMARQUE. De ce qui a été dit plus haut il s'ensuit que si le domaine $\mathscr{K}$ de solutions d'un système normal ne contient aucun sommet, ce domaine est vide, et le système n'a pas donc de solutions (il est incompatible).

● EXEMPLE 1. Trouver tous les sommets du domaine $\mathscr{K}$ défini par le système d'inégalités suivant

$$\left.\begin{aligned} x + y + 1 &\geqslant 0, \\ x - 2y - 2 &\geqslant 0, \\ 2x - y - 4 &\geqslant 0. \end{aligned}\right\}$$

En résolvant les sous-systèmes

$$\left.\begin{aligned} x + y + 1 &= 0, \\ x - 2y - 2 &= 0, \end{aligned}\right\} \quad \left.\begin{aligned} x + y + 1 &= 0, \\ 2x - y - 4 &= 0, \end{aligned}\right\}$$

$$\left.\begin{aligned} x - 2y - 2 &= 0, \\ 2x - y - 4 &= 0 \end{aligned}\right\}$$

(tous les trois étant réguliers), on trouve trois points:

$$(0, -1), \quad (1, -2), \quad (2, 0),$$

dont seuls les deux derniers satisfont à toutes les inégalités données. Le domaine $\mathscr{K}$ a donc pour sommets les points

$$A_1 (1, -2) \quad \text{et} \quad A_2 (2, 0).$$

Revenons au système (1). Soient

$$A_1, A_2, \ldots, A_p$$

tous les sommets du domaine $\mathscr{K}$. L'ensemble $\langle A_1, A_2, \ldots, A_p \rangle$, enveloppe convexe du système de points A_1, $A_2 \ldots, A_p$, appartient lui aussi à $\mathscr{K}$ ($\mathscr{K}$ étant un domaine

convexe!). Mais dans ce cas, selon le lemme 1, l'ensemble

$$\langle A_1, A_2, \ldots, A_p\rangle + K_0$$

appartient également à $\mathscr{K}$. Nous allons montrer maintenant qu'en réalité cette somme se confond avec $\mathscr{K}$, c'est-à-dire a lieu le théorème suivant.

● THÉORÈME.. *Si un système d'inégalités est normal, on a*

$$\mathscr{K} = \langle A_1, A_2, \ldots, A_p\rangle + \mathscr{K}_0, \tag{7}$$

où $A_1, A_2, \ldots, A_p$ *sont tous les sommets du domaine* $\mathscr{K}$.

● DÉMONSTRATION. Soit P un point arbitraire du domaine $\mathscr{K}$ différant de ses sommets. La droite A_1P coupe le domaine convexe $\mathscr{K}$ ou bien suivant un segment A_1A (fig. 34), ou bien suivant un rayon dont l'origine est A_1 (fig. 35). Dans ce deuxième cas $P - A_1 \in \mathscr{K}_0$ (lemme 2), par conséquent, $P \in A_1 + \mathscr{K}_0$. Dans le premier cas nous raisonnons comme suit : si le point A se trouve sur une arête limitée A_iA_j du domaine $\mathscr{K}$ (cf. fig. 34), le point P appartient alors à l'enveloppe convexe des points A_1, A_i, A_j, si A se trouve sur une arête illimitée dont l'origine se confond avec le sommet A_i (fig. 36), on a, d'après le lemme 1, $A \in \in A_i + \mathscr{K}_0$, d'où $P \in \langle A_1, A_i\rangle + \mathscr{K}_0$. De cette façon, dans tous les cas le point P appartient à l'ensemble $\langle A_1, A_2, \ldots \ldots, A_p\rangle + \mathscr{K}_0$. Le théorème est donc démontré.

Vu que le procédé de détermination des sommets nous est déjà familier, il nous reste donc à apprendre à déterminer le domaine $\mathscr{K}_0$ pour pouvoir décrire entièrement le domaine $\mathscr{K}$. Le domaine $\mathscr{K}_0$ représente le domaine de solutions d'un système normal homogène (2). Nous passons à son examen.

C. *Système normal homogène d'inégalités* (2). Chacune des inégalités (2) définit un demi-plan dont la droite limite passe par l'origine des coordonnées. La partie commune de ces demi-plans est justement le domaine $\mathscr{K}_0$.

Fig. 34

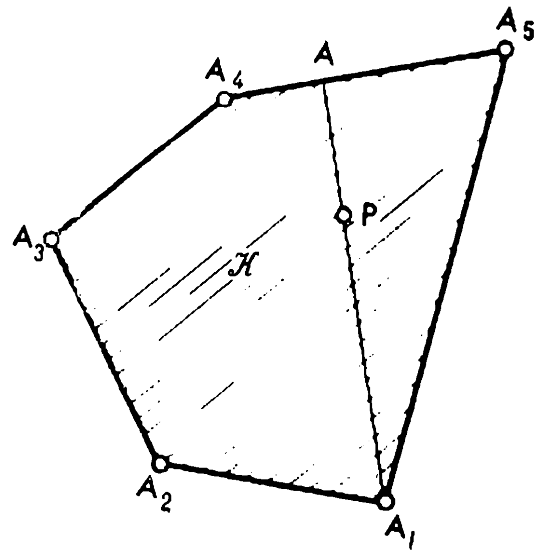
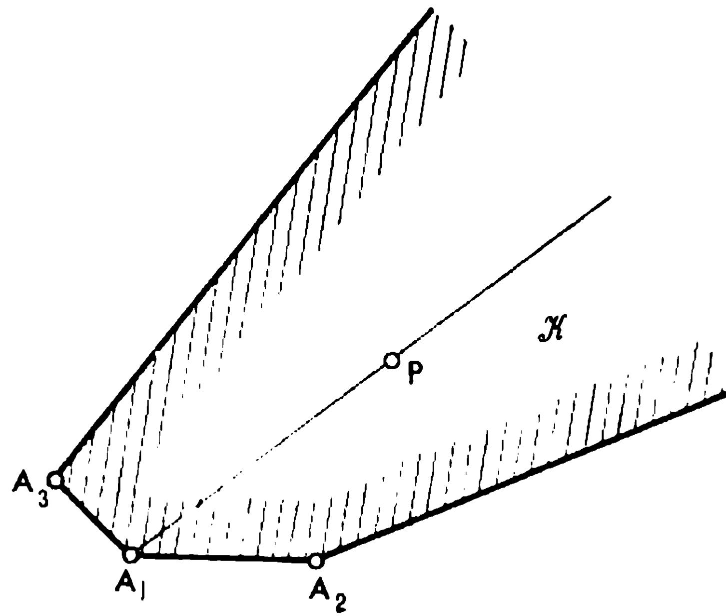


Fig. 35



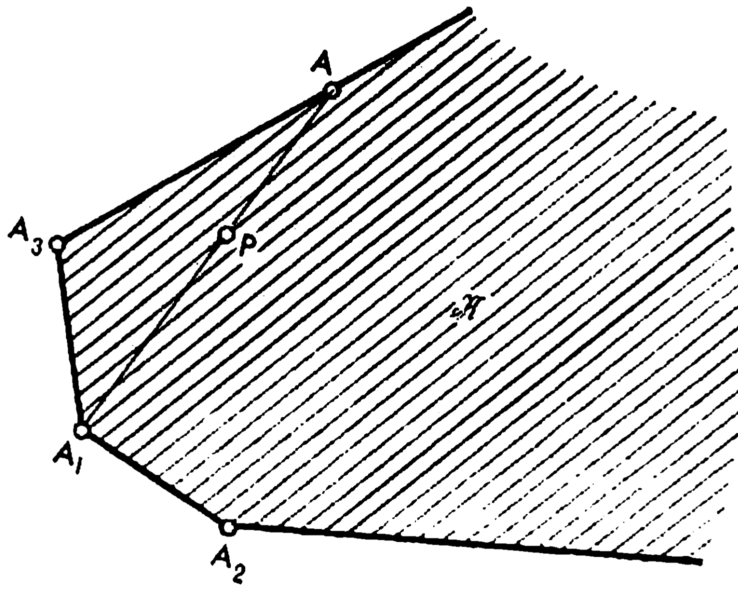


Fig. 36

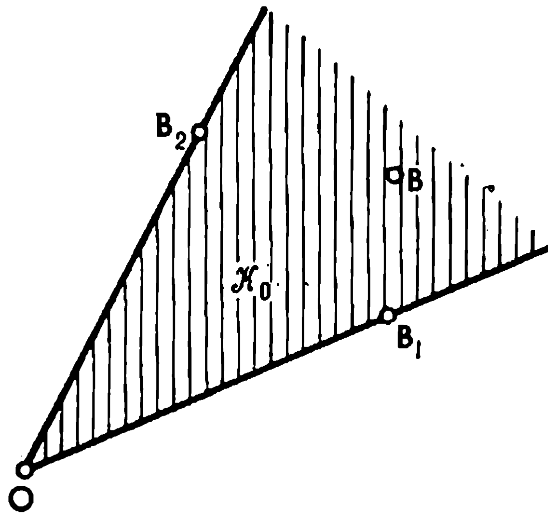


Fig. 37

Dans le cas considéré, parmi les droites limites il y en a au moins deux qui diffèrent (le système (2) étant normal !). Par conséquent, $\mathscr{K}_0$ ou bien coïncide avec l'origine des coordonnées ($x = 0$, $y = 0$), ou bien représente un rayon issu de l'origine des coordonnées, ou, enfin, est un certain angle, inférieur à 180°, dont le sommet est à l'origine des coordonnées. Si l'on connaît deux points B_1 et B_2 situés sur les différents côtés de cet angle (fig. 37), tous les points de ce dernier prennent alors la forme

$$B = t_1 B_1 + t_2 B_2, \tag{8}$$

où t_1 et t_2 sont des nombres non négatifs quelconques. Il est facile de déterminer les points B_1 et B_2 si l'on tient compte que chacun de ces points : a) appartient à $\mathscr{K}_0$, c'est-à-dire vérifie le système (2), et b) se trouve sur la frontière de $\mathscr{K}_0$, c'est-à-dire satisfait à l'une des équations (3). Si, par contre, $\mathscr{K}_0$ est un rayon, au lieu de (8) on aura

$$B = tB_1, \tag{9}$$

où B_1 est un point quelconque de ce rayon (différant de l'origine des coordonnées), t un nombre non négatif quelconque.

● EXEMPLE 2. Trouver le domaine $\mathscr{K}_0$ de solutions du système

$$\left.\begin{aligned} x + y &\geqslant 0, \\ x - 2y &\geqslant 0, \\ 2x - y &\geqslant 0, \end{aligned}\right\} \tag{10}$$

ainsi que le domaine $\mathscr{K}$ de solutions du système de l'exemple 1.

● SOLUTION. Le système (10) est normal : l'unique solution du système d'équations homogène correspondant

$$\left.\begin{aligned} x + y &= 0, \\ x - 2y &= 0, \\ 2x - y &= 0 \end{aligned}\right\} \tag{11}$$

est donc (0, 0).

Choisissons maintenant un point quelconque vérifiant la première des équations (11) (différant de (0, 0)), C (—1, 1), par exemple. On s'assure par une vérification que le point C ne vérifie pas toutes les inégalités (10), par conséquent, ni C, ni aucun autre point du rayon OC (différant de l'origine O) n'appartiennent pas à $\mathscr{K}_0$. Après avoir examiné le point $-C$ (c'est-à-dire le point (1, —1)), nous trouvons qu'il appartient à $\mathscr{K}_0$. Ainsi, $B_1 = (1, -1)$. Le point (2, 1) satisfait à la deuxième équation du système considéré ; il représente donc lui aussi une solution du système (10), par conséquent, $B_2 = (2, 1)$. Le domaine $\mathscr{K}_0$ se compose de points (fig. 38)

$$t_1B_1 + t_2B_2 = t_1(1, -1) + t_2(2, 1) = (t_1 + 2t_2, -t_1 + t_2),$$

où t_1 et t_2 sont des nombres non négatifs quelconques.

En revenant au système d'inégalités de l'exemple 1, nous voyons que le système d'inégalités homogène qui lui correspond est justement (10). D'après le théorème démontré ci-dessus on a

$$\mathscr{K} = \langle A_1, A_2\rangle + \mathscr{K}_0,$$

où A_1 (1, —2) et A_2 (2, 0) sont les sommets du domaine $\mathscr{K}$. Ainsi, $\mathscr{K}$ est constitué des points (fig. 39)

$$s(1, -2) + (1 - s)(2, 0) + (t_1 + 2t_2, -t_1 + t_2) = (2 - s + t_1 + 2t_2, -2s - t_1 + t_2),$$

s étant un nombre arbitraire de l'intervalle [0, 1], t_1t_2 les nombres non négatifs quelconques.

● EXEMPLE 3. Trouver le domaine de solutions du système

$$\left.\begin{aligned} 2x - y &\geqslant 0,\\ -4x + 2y &\geqslant 0,\\ x + y &\geqslant 0. \end{aligned}\right\}$$

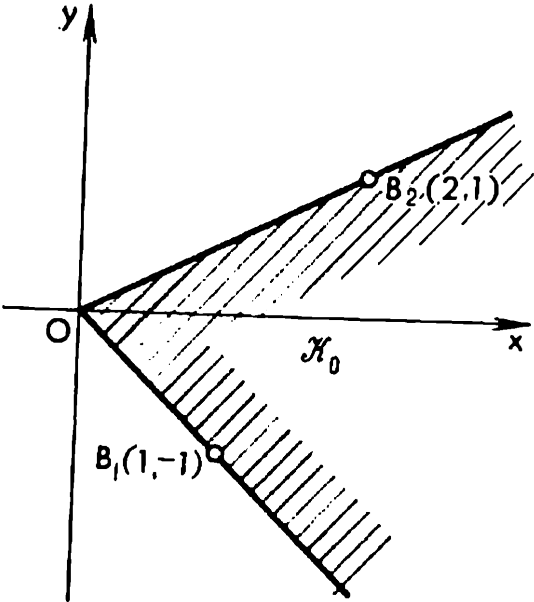


Fig. 38

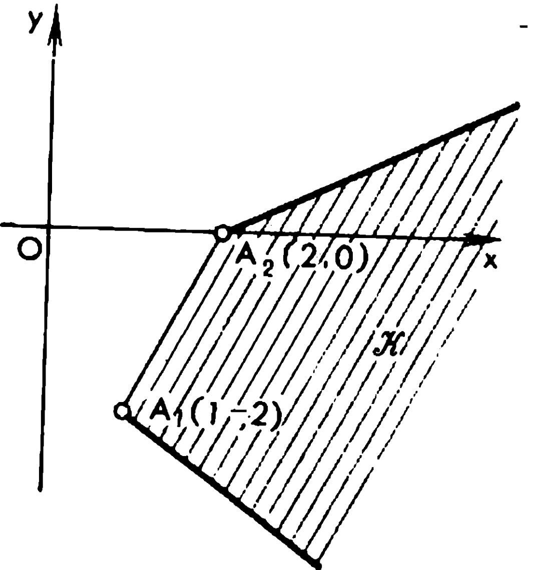


Fig. 39

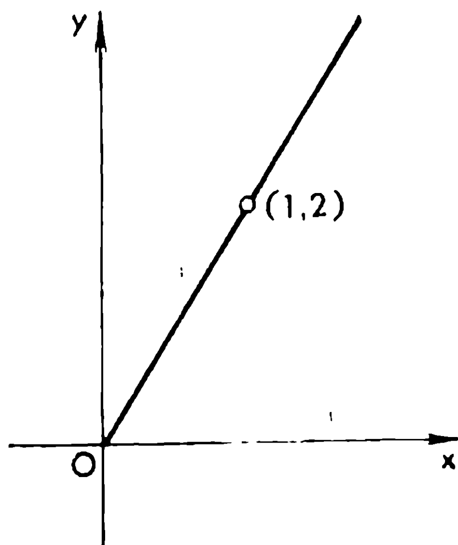


Fig. 40

En procédant comme dans l'exemple 2, on ne trouve qu'un seul rayon :

$$B = t\,(1,\ 2) = (t,\ 2t), \quad (t \geqslant 0)$$

(fig. 40).

● EXEMPLE 4. Trouver le domaine de solutions du système

$$\left.\begin{aligned} 2x - y &= 0, \\ x + y &= 0, \\ -3x + y &\geqslant 0. \end{aligned}\right\}$$

Dans ce cas aucune des équations

$$\begin{aligned} 2x - y &= 0, \\ x + y &= 0, \\ -3x + y &= 0 \end{aligned}$$

ne possède de solutions (excepté (0, 0)) qui vérifieraient toutes les inégalités données. Le domaine $\mathscr{K}_0$ ne contient qu'un seul point (0, 0), l'origine des coordonnées.

D. *Cas où le système d'inégalités* (1) *n'est pas normal.* Cela signifie que le domaine de solutions $\mathscr{L}$ d'un système homogène d'équations (3) contient non seulement l'origine des coordonnées. Par conséquent, toutes les équations (3)

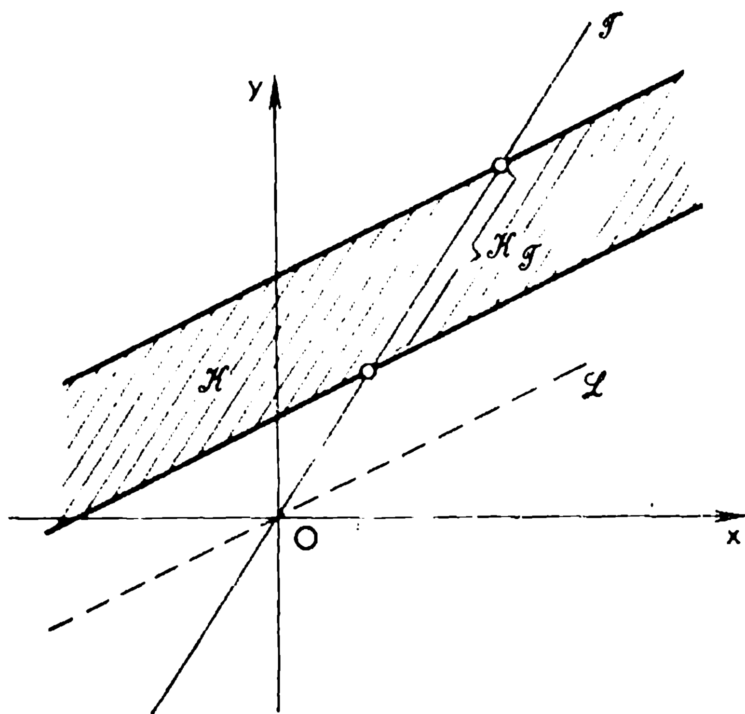


Fig. 41

définissent dans le plan une seule et même droite, notamment $\mathscr{L}$.

En vertu du lemme 1 le domaine $\mathscr{K}$ contient, en plus des points P, la droite $P + \mathscr{L}$ (passant par P parallèlement à $\mathscr{L}$). Considérons une droite quelconque $\mathscr{T}$ non parallèle à $\mathscr{L}$. Si l'on connaît les points de la droite $\mathscr{T}$ appartenant au domaine $\mathscr{K}$, nous désignerons cet ensemble par $\mathscr{K}_{\mathscr{T}}$, on pourra définir le domaine $\mathscr{K}$, car dans ce cas $\mathscr{K} = \mathscr{K}_{\mathscr{T}} + \mathscr{L}$ (fig. 41).

L'équation de la droite $\mathscr{L}$ est $a_1x + b_1y = 0$. Dans cette équation l'un des coefficients a_1 ou b_1 est non nul; soit, par exemple, $b_1 \neq 0$. Dans ce cas-là, on peut prendre pour la droite $\mathscr{T}$, non parallèle à $\mathscr{L}$, l'axe y (son équation est $x = 0$). L'ensemble $\mathscr{K}_{\mathscr{T}}$ désigné cette fois par $\mathscr{K}_y$ représentera alors une partie de l'axe y qui se trouve dans $\mathscr{K}$. Pour trouver cet ensemble, il convient de poser $x = 0$ dans le système (1). On obtient alors le système d'iné-

galités

$$\left.\begin{array}{r} b_1y + c_1 \geqslant 0, \\ b_2y + c_2 \geqslant 0, \\ \dots\dots\dots \\ b_my + c_m \geqslant 0 \end{array}\right\} \qquad (12)$$

à une seule inconnue y, bien facile à résoud e *). Il tient à remarquer que l'ensemble $\mathscr{K}_y$ peut représenter ou bien un ensemble vide ($\mathscr{K}$ sera alors aussi vide), ou bien un point, un segment, un rayon (mais non pas l'axe y tout entier, car autrement $\mathscr{K}$ représentera tout un plan, ce qui est impossible). En déterminant cet ensemble, on définit en même temps le domaine $\mathscr{K}$, car

$$\mathscr{K} = \mathscr{K}_y + \mathscr{L} \qquad (13)$$

(si $\mathscr{L}$ n'est pas parallèle à l'axe y).

● EXEMPLE 5. Trouver le domaine $\mathscr{K}$ de solutions du système

$$\left.\begin{array}{r} x + y - 1 \geqslant 0, \\ -x - y + 2 \geqslant 0, \\ 2x + 2y + 3 \geqslant 0. \end{array}\right\}$$

Il est aisé de voir que le système ci-dessus n'est pas normal et $\mathscr{L}$ est une droite

$$x + y = 0$$

(non parallèle à l'axe y). En y posant $x = 0$, on obtient

$$\left.\begin{array}{r} y - 1 \geqslant 0, \\ -y + 2 \geqslant 0, \\ 2y + 3 \geqslant 0, \end{array}\right\}$$

*) Il est à remarquer que le système (12) (considéré en tant qu'un système d'inégalités à une seule inconnue) sera déjà normal. En effet, dans le cas contraire le système homogène qui lui correspond aurait une solution non nulle; mais alors le système (3) aurait aussi une solution différente de (0, 0).

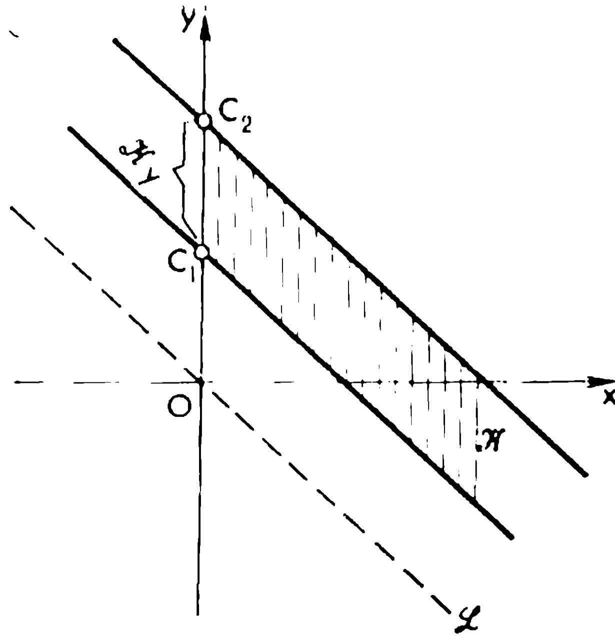


Fig. 42

d'où l'on comprend que $\mathcal{K}_y$, intersection de $\mathcal{K}$ avec l'axe y, est un segment aux extrémités C_1 (0, 1) et C_2 (0, 2). $\mathcal{K}$ représente donc un ensemble des points de la forme (fig. 42)

$$(0,\ y) + (x,\ -x) = (x,\ y - x),$$

où x est un nombre arbitraire et y un nombre quelconque dans l'intervalle de 1 à 2.

Pour conclure arrêtons-nous encore sur un théorème qui résulte des déductions faites plus haut. Dans le cas bidimensionnel considéré ici (quand tout se passe dans un plan) ce théorème ne paraît pas tellement important et il serait plus juste de le considérer comme point de départ de son extension au cas « à n dimensions » (voir § 7).

● THÉORÈME. *Tout domaine polygonal convexe (non vide) $\mathcal{K}$ dans un plan peut être mis sous la forme d'une somme*

$$\langle A_1,\ A_2,\ \ldots,\ A_p\rangle + (B_1,\ B_2,\ \ldots,\ B_q). \qquad (14)$$

Le premier terme de cette somme est l'enveloppe convexe d'un système de points $A_1, A_2, \ldots, A_p$, le deuxième, un ensemble des points de la forme $t_1B_1 + t_2B_2 + \ldots + t_qB_q$, où $t_1, t_2, \ldots, t_q$ sont des nombres négatifs quelconques.

La démonstration de ce théorème peut s'effectuer en quelques mots. Considérons un système d'inégalités qui définit le domaine $\mathcal{K}$.

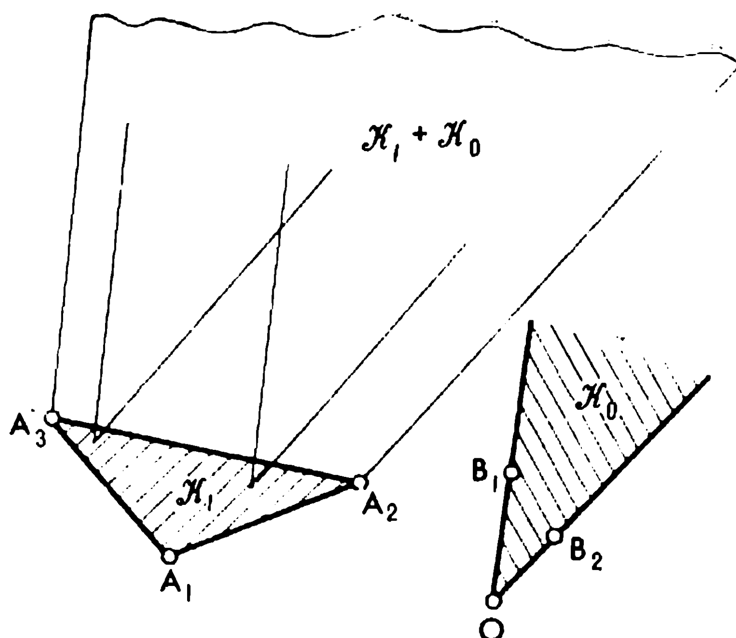


Fig. 43

S'il est normal, on a l'égalité (7) ; comme dans cette égalité $\mathcal{K}_0$ est l'un des ensembles de la forme (B_1, B_2), (B_1) ou (O) (l'origine des coordonnées), on est amené à conclure que notre assertion est juste dans le cas d'un système normal. Si, par contre, ce système n'est pas normal, on obtient l'égalité (13) qui permet également de mettre $\mathcal{K}$ sous la forme requise (pourquoi?).

Il est à remarquer que si tous les points $A_1, A_2, \ldots, A_p$ coïncident avec l'origine des coordonnées O, l'ensemble $\langle A_1, A_2, \ldots, A_p\rangle$ se confond lui aussi avec O ; il ne reste alors de la somme (14) que le deuxième terme. Dans le cas où les points $B_1, B_2, \ldots, B_q$ coïncident avec O, l'ensemble $(B_1, B_2, \ldots, B_q)$ l'est aussi et la somme (14) n'a que son premier terme.

La réciproque a également lieu bien qu'avec une certaine restriction :

● THÉORÈME. *Tout ensemble de la forme*

$$\langle A_1, A_2, \ldots, A_p\rangle + (B_1, B_2, \ldots, B_q)$$

représente dans un plan soit un plan entier soit un domaine polygonal convexe dans celui-ci.

La démonstration est évidente. Le deuxième terme, c'est-à-dire le domaine $\mathcal{K}_0 = (B_1, B_2, \ldots, B_q)$ est soit un plan entier, soit un emi-plan, soit un angle (inférieur à 180°), soit un rayon, soit, enfin, un point (origine des coordonnées). Le premier terme $\mathcal{K}_1 = \langle A_1, A_2, \ldots, A_p\rangle$ représente un polygone convexe. L'ensemble

$\mathcal{K}_1 + \mathcal{K}_2$ s'obtient en effectuant des translations de $\mathcal{K}_0$ à la longueur des segments OK_1 (où K_1 est un nombre quelconque de $\mathcal{K}_1$) et en prenant ensuite la réunion des ensembles obtenus (fig. 43). Il est aisé de voir que l'on obtient ainsi soit un plan entier (il en sera ainsi si $\mathcal{K}_0$ est le plan entier), soit un domaine polygonal convexe dans celui-ci.

§ 6. DOMAINE DE SOLUTIONS D'UN SYSTÈME À TROIS INCONNUES

Après l'examen détaillé fait au paragraphe précédent, nous pouvons maintenant réduire au minimum l'étude théorique du système à trois inconnues.

En plus du système initial

$$\left.\begin{array}{l} a_1x + b_1y + c_1z + d_1 \geqslant 0, \\ \dots\dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz + d_m \geqslant 0 \end{array}\right\} \qquad (1)$$

nous allons considérer de nouveau (comme au § 5) deux systèmes :

$$\left.\begin{array}{l} a_1x + b_1y + c_1z \geqslant 0, \\ \dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz \geqslant 0 \end{array}\right\} \qquad (2)$$

et

$$\left.\begin{array}{l} a_1x + b_1y + c_1z = 0, \\ \dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz = 0. \end{array}\right\} \qquad (3)$$

Désignons par $\mathcal{K}$ le domaine de solutions du système (1), par $\mathcal{K}_0$ celui du système (2) et par $\mathcal{L}$ celui du système (3). En faisant recours à la terminologie introduite plus haut, on peut dire que $\mathcal{K}$ est un domaine polyèdre convexe dans l'espace et $\mathcal{K}_0$ un cône polyèdre convexe. Les lemmes 1 et 2 du § 5 restent, comme nous l'avons déjà signalé, valides dans ce cas aussi.

A. *Cas où le système d'inégalités* (1) *est normal.* Dans ce cas le domaine $\mathscr{K}$ ne renferme pas de droites et a, par conséquent, au moins un sommet. En effet, si $\mathscr{K}$ est situé dans un plan (ceci est d'ailleurs bien possible, comme nous l'avons dit dans le § 2), $\mathscr{K}$ représente alors un domaine polygonal convexe dans un plan ne contenant pas de droites et doit (voir B du § 5) nécessairement avoir des sommets. Si, par contre, le domaine $\mathscr{K}$ n'est pas situé dans un plan, nous considérerons sa frontière. Cette dernière se compose de faces planes dont chacune, en tant qu'un domaine polygonal convexe ne contenant pas de droites, doit avoir des sommets. Mais il est aisé de voir que le sommet de n'importe quelle face l'est à la fois du domaine $\mathscr{K}$.

Chaque sommet A du domaine $\mathscr{K}$ est le point où convergent au moins trois plans limites pour lesquels le point A est le point commun *unique*. En effet, s'il était autrement, tous les plans limites passant par A ou bien coïncideraient, ou bien auraient une droite commune. Mais alors le segment suffisamment petit passant par A et se trouvant dans le plan limite commun ou sur la droite limite commune appartiendrait à $\mathscr{K}$, ce qui contredit la définition du sommet.

Comme il résulte de l'exposé ci-dessus, nous devons faire des changements dans le procédé de détermination des sommets (voir B du § 5). Notamment, nous entendrons par *sous-système régulier* non pas un sous-système avec deux, mais avec *trois* équations du système

$$\left.\begin{array}{r} a_1x + b_1y + c_1z + d_1 = 0, \\ \dots\dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz + d_m = 0 \end{array}\right\} \qquad (4)$$

à condition que ce sous-système n'ait qu'une seule solution (x, y, z). Pour cette définition du sous-système régulier le procédé de détermination des sommets reste exactement le même qu'auparavant, à savoir :

Pour déterminer tous les sommets du domaine $\mathscr{K}$ il convient de trouver les solutions de tous les sous-systèmes réguliers du système (4) *pour en choisir ensuite celles qui vérifient le système initial* (1).

Le théorème du numéro B du § 5 reste évidemment en vigueur; les modifications qu'il faut apporter à la démonstration de celle-ci sont évidentes. L'assertion suivant laquelle un système normal n'a pas de solutions si le domaine $\mathscr{K}$ est sans sommets reste également vraie.

● EXEMPLE 1. Trouver les sommets du domaine $\mathscr{K}$ défini par le système d'inégalités

$$\left.\begin{aligned} 2x + y + z - 1 &\geqslant 0,\\ x + 2y + z - 1 &\geqslant 0,\\ x + y + 2z - 1 &\geqslant 0,\\ x + y + z - 1 &\geqslant 0. \end{aligned}\right\} \qquad (5)$$

Dans le cas en question le système d'équations homogène correspondant a pour expression

$$\left.\begin{aligned} 2x + y + z &= 0,\\ x + 2y + z &= 0,\\ x + y + 2z &= 0,\\ x + y + z &= 0. \end{aligned}\right\}$$

En résolvant ce système, on se convainc que la seule solution possible est (0, 0, 0); c'est-à-dire que le système (5) est normal.

Pour trouver les sommets, il nous faut examiner les différents sous-systèmes de trois équations du système (4):

$$\left.\begin{aligned} 2x + y + z - 1 &= 0,\\ x + 2y + z - 1 &= 0,\\ x + y + 2z - 1 &= 0; \end{aligned}\right\} \qquad \left.\begin{aligned} 2x + y + z - 1 &= 0,\\ x + 2y + z - 1 &= 0,\\ x + y + z - 1 &= 0; \end{aligned}\right\}$$

$$\left.\begin{aligned} 2x + y + z - 1 &= 0,\\ x + y + 2z - 1 &= 0,\\ x + y + z - 1 &= 0; \end{aligned}\right\} \qquad \left.\begin{aligned} x + 2y + z - 1 &= 0,\\ x + y + 2z - 1 &= 0,\\ x + y + z - 1 &= 0. \end{aligned}\right\}$$

Après avoir effectué les calculs nécessaires, on trouve que tous les sous-systèmes sont réguliers et ont pour leurs solutions les points

$$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right), \ (0, 0, 1), \ (0, 1, 0), \ (1, 0, 0),$$

dont le premier ne vérifie pas le système (5) et les trois autres le font. Par conséquent, le domaine $\mathcal{K}$ a pour ses sommets:

$$A_1 (1, 0, 0), \ A_2 (0, 1, 0), \ A_3 (0, 0, 1).$$

B. *Système homogène normal d'inégalités* (2). Chacune des inégalités (2) définit un demi-espace dont le plan limite passe par l'origine des coordonnées.

Dans le cas considéré les plans limites se coupent en un seul point, l'origine des coordonnées (le système (2) est donc normal!). En d'autres termes, l'ensemble $\mathcal{K}_0$, domaine de solutions du système (2), est un cône polyèdre convexe à un *seul* sommet. De l'énumération des cônes polyèdres convexes donnée au § 4 s'ensuit que dans le cas considéré $\mathcal{K}_0$ est soit une pyramide infinie convexe, soit un angle plan, soit un rayon, soit enfin un point (l'origine des coordonnées). Laissons pour le moment ce dernier cas de côté. Dans tous les autres cas on a

$$\mathcal{K}_0 = (B_1, B_2, \ldots, B_q),$$

où $B_1, B_2, \ldots, B_q$ sont des points quelconques dont chacun se trouve sur chaque arête du cône $\mathcal{K}_0$ (cf. le théorème 2 du § 4). Pour trouver ces points on part du raisonnement suivant. Chacun de ceux-ci: a) appartient à $\mathcal{K}_0$, c'est-à-dire satisfait au système (2) et b) appartient à l'intersection

de deux faces différentes, c'est-à-dire vérifie deux équations non proportionnelles *) du système (3).

S'il s'avère que le seul point vérifiant les conditions a) et b) est celui (0, 0, 0), le domaine $\mathscr{K}_0$ se confond alors avec l'origine des coordonnées.

● EXEMPLE 2. Trouver le domaine $\mathscr{K}_0$ de solutions du système

$$\left.\begin{aligned} 2x + y + z &\geqslant 0, \\ x + 2y + z &\geqslant 0, \\ x + y + 2z &\geqslant 0, \\ x + y + z &\geqslant 0, \end{aligned}\right\} \tag{6}$$

de même que le domaine $\mathscr{K}$ de solutions du système de l'exemple 1.

Il est à remarquer avant tout que le système (6) est lié au système d'inégalités (5) de l'exemple 1 ; à savoir : (6) est le système homogène correspondant à (5). Par suite, le système (6) est normal.

Dans le cas en question un système de deux équations non proportionnelles peut être représenté de six manières différentes :

$$\left.\begin{aligned} x + 2y + z &= 0, \\ x + y + 2z &= 0; \end{aligned}\right\} \quad \left.\begin{aligned} 2x + y + z &= 0, \\ x + y + 2z &= 0; \end{aligned}\right\}$$

$$\left.\begin{aligned} 2x + y + z &= 0, \\ x + y + z &= 0; \end{aligned}\right\} \quad \left.\begin{aligned} 2x + y + z &= 0, \\ x + 2y + z &= 0; \end{aligned}\right\}$$

$$\left.\begin{aligned} x + 2y + z &= 0, \\ x + y + z &= 0; \end{aligned}\right\} \quad \left.\begin{aligned} x + y + 2z &= 0, \\ x + y + z &= 0. \end{aligned}\right\}$$

---

*) Les équations $ax + by + cz = 0$ et $a'x + b'y + c'z = 0$ sont dites « non proportionnelles » si au moins l'une des égalités $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$ n'est pas satisfaite ; dans ce cas les plans correspondants se coupent suivant une droite quelconque.

Choisissons pour chacun de ces six systèmes deux solutions non nulles: (x, y, z) et $(-x, -y, -z)$. Ainsi, par exemple, pour le premier système on peut prendre $(3, -1, -1)$ et $(-3, 1, 1)$; les inégalités (6) ne sont vérifiées que pour la première de ces solutions. D'où l'on obtient le point $B_1 = (3, -1, -1)$. En procédant de façon analogue avec les autres cinq systèmes, on obtient les points $B_2 = (-1, 3, -1)$ et $B_3 = (-1, -1, 3)$. Ainsi, le domaine $\mathcal{K}_0$ se compose des points de la forme

$$t_1B_1 + t_2B_2 + t_3B_3 =$$
$$= (3t_1 - t_2 - t_3, \ -t_1 + 3t_2 - t_3, \ -t_1 - t_2 - 3t_3),$$

où t_1, t_2, t_3 sont des nombres non négatifs quelconques.

Revenons maintenant au système d'inégalités (5) de l'exemple 1. Le système homogène qui lui correspond est, comme on l'a déjà dit celui (6). Par conséquent, le domaine $\mathcal{K}$ a pour expression

$$\langle A_1, A_2, A_3\rangle + \mathcal{K}_0$$

et se compose des points

$$s_1A_1 + s_2A_2 + s_3A_3 + t_1B_1 + t_2B_2 + t_3B_3 =$$
$$= s_1 (1, 0, 0) + s_2 (0, 1, 0) + s_3 (0, 0, 1) +$$
$$+t_1 (3, -1, -1) + t_2 (-1, 3, -1) + t_3 (-1, -1, 3) =$$
$$= (s_1 + 3t_1 - t_2 - t_3, \ s_2 - t_1 + 3t_2 - t_3,$$
$$s_3 - t_1 - t_2 + 3t_3),$$

où t_1, t_2, t_3 sont des nombres non négatifs quelconques et s_1, s_2, s_3 sont également des nombres non négatifs dont la somme vaut 1.

C. *Cas où le système d'inégalités* (1) *n'est pas normal.* Cela signifie que le domaine $\mathcal{L}$ de solutions du système homogène d'équations (3) contient des points différant de l'origine

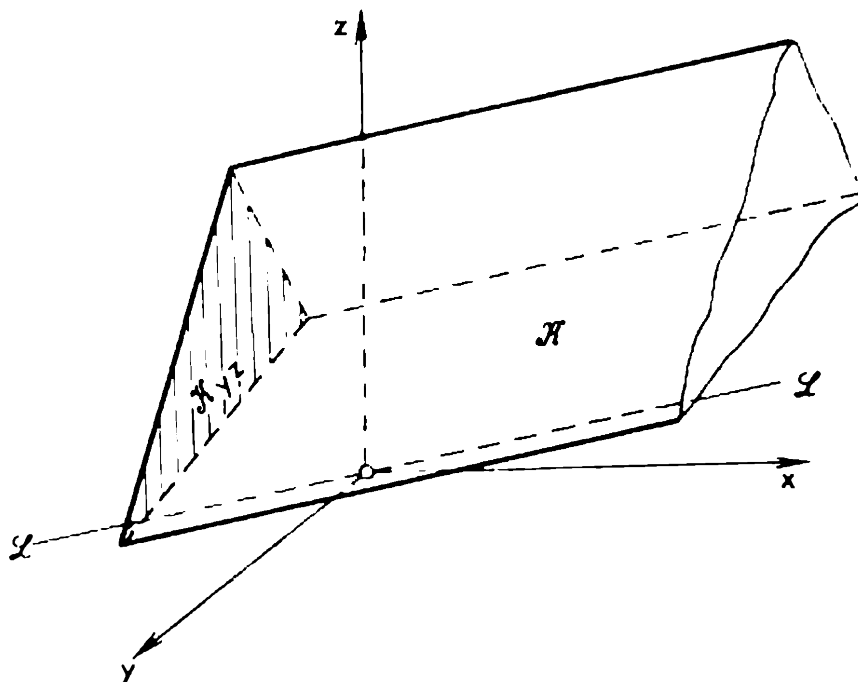


Fig. 44

des coordonnées. Comme $\mathcal{L}$ est une intersection des plans, les deux cas suivants peuvent alors se présenter:

1. $\mathcal{L}$ est une droite. En vertu du lemme 1 le domaine $\mathcal{K}$ renferme avec tout son point P la droite $P + \mathcal{L}$. Considérons maintenant un plan $\mathcal{T}$ quelconque non parallèle à $\mathcal{L}$. Si l'on connaît les points du plan $\mathcal{T}$ appartenant au domaine $\mathcal{K}$, dont l'ensemble sera désigné par $\mathcal{K}_{\mathcal{T}}$, on peut alors définir le domaine $\mathcal{K}$ lui-même, car dans ce cas $\mathcal{K} = \mathcal{K}_{\mathcal{T}} + \mathcal{L}$.

Mais, quelle que soit la droite $\mathcal{L}$, on peut toujours prendre pour un plan $\mathcal{T}$ non parallèle à celle-ci l'un des plans des coordonnées xOy, xOz ou yOz. Admettons, par exemple, que $\mathcal{L}$ n'est pas parallèle au plan yOz. Prenons donc ce dernier pour le plan $\mathcal{T}$. Dans ce cas l'ensemble $\mathcal{K}_{\mathcal{T}}$, qui sera désigné par la suite par $\mathcal{K}_{y,\,z}$, est la partie du plan yOz comprise dans $\mathcal{K}$ (fig. 44). Pour trouver cet ensemble il convient de poser $x = 0$ dans le système (1). On obtient alors le

système d'inégalités *)

$$\left.\begin{array}{r} b_1y + c_1z + d_1 \geqslant 0, \\ \dots\dots\dots\dots \\ b_my + c_mz + d_m \geqslant 0, \end{array}\right\} \qquad (7)$$

que l'on peut résoudre par la méthode exposée au § 5.

Après avoir trouvé l'ensemble $\mathcal{K}_{y,z}$, on peut écrire

$$\mathcal{K} = \mathcal{K}_{y,z} + \mathcal{L} \qquad (8)$$

(si la droite $\mathcal{L}$ n'est pas parallèle au plan yOz). Ainsi, le domaine $\mathcal{K}$ est entièrement décrit.

● REMARQUE.. S'il se trouve que l'ensemble $\mathcal{K}_{y,z}$ est vide, $\mathcal{K}$ le sera aussi. Cela montre que le système (1) est incompatible.

● EXEMPLE 3. Définir le domaine $\mathcal{K}$ de solutions du système

$$\left.\begin{array}{r} -2x + y + z - 1 \geqslant 0, \\ -3x - y + 4z - 1 \geqslant 0, \\ -x - 2y + 3z \geqslant 0. \end{array}\right\} \qquad (9)$$

Examinons un système d'équations homogène correspondant

$$\left.\begin{array}{r} -2x + y + z = 0, \\ -3x - y + 4z = 0, \\ -x - 2y + 3z = 0. \end{array}\right\} \qquad (10)$$

En résolvant ce système, on remarque aussitôt que la troisième équation se déduit des deux premières. Le système se ramène donc à deux premières équations. L'ensemble $\mathcal{L}$ de ses solutions est la droite suivant laquelle se coupent

*) Le système (7) comme on le voit aisément, sera cette fois normal (cf. la note analogue à la page 354).

les plans

$$-2x + y + z = 0$$

et

$$-3x - y + 4z = 0.$$

Choisissons sur la droite $\mathscr{L}$ un point quelconque B différant de l'origine des coordonnées. Pour cela il suffit de trouver trois nombres x, y, z quelconques (non simultanément nuls) vérifiant les deux premières équations du système (10). Prenons, par exemple, 1, 1, 1. Ainsi, $\mathscr{L}$ est la droite OB, où $B = (1, 1, 1)$.

Nous voyons que la droite $\mathscr{L}$ n'est pas parallèle, par exemple, au plan des coordonnées yOz. En posant ensuite $x = 0$ dans (9) on obtient le système

$$\left.\begin{aligned} y + z - 1 &\geqslant 0,\\ -y + 4z - 1 &\geqslant 0,\\ -2y + 3z &\geqslant 0 \end{aligned}\right\}$$

à deux inconnues y et z qui est évidemment normal. La méthode exposée au § 5 permet de définir le domaine $\mathscr{K}_{y,z}$ de ses solutions. Après avoir effectué les calculs nécessaires, on trouve que $\mathscr{K}_{y,z}$ est un ensemble qui se compose d'un seul point $A\left(\frac{3}{5}, \frac{2}{5}\right)$ (dans le plan yOz).

Par conséquent, le domaine cherché $\mathscr{K}$ est constitué de tous les points de la forme

$$A + tB = \left(0, \frac{2}{5}, \frac{3}{5}\right) + t(1, 1, 1) = \left(t, \frac{3}{5} + t, \frac{2}{5} + t\right),$$

où t est un nombre non négatif quelconque (le domaine $\mathscr{K}$ est une droite parallèle à $\mathscr{L}$.

2. $\mathscr{L}$ est un plan. On prend alors pour l'ensemble sécant $\mathscr{T}$ une droite quelconque non parallèle à ce plan, l'un des axes des coordonnées, par exemple. Admettons que l'axe z ne soit pas parallèle à $\mathscr{L}$; prenons-le pour $\mathscr{T}$. Pour trouver

Fig. 45

l'ensemble $\mathscr{K}_z$, partie de l'axe z comprise dans $\mathscr{K}$, il faut poser $x = 0$, $y = 0$ dans le système (1). On obtient alors le système d'inégalités

$$\left.\begin{array}{r} c_1 z + d_1 \geqslant 0, \\ \dots\dots\dots\dots \\ c_m z + d_m \geqslant 0, \end{array}\right\} \qquad (11)$$

bien facile à résoudre *). Après avoir trouvé l'ensemble $\mathscr{K}_z$, on peut écrire (fig. 45)

$$\mathscr{K} = \mathscr{K}_z + \mathscr{L} \qquad (12)$$

(si le plan $\mathscr{L}$ n'est pas parallèle à l'axe z). Le domaine $\mathscr{K}$ est donc entièrement décrit.

---

*) Le système (11) est normal.

● REMARQUE. Si l'ensemble $\mathcal{K}_z$ s'avère vide, $\mathcal{K}$ le sera aussi. Dans ce cas le système (1) est incompatible.

● EXEMPLE 4. Trouver le domaine $\mathcal{K}$ de solutions du système

$$\left.\begin{aligned} x - y + z + 1 \geqslant 0, \\ -x + y - z + 2 \geqslant 0. \end{aligned}\right\} \tag{13}$$

Dans le cas considéré le système d'équations homogène correspondant a pour expression

$$\left.\begin{aligned} x - y + z = 0, \\ -x + y - z = 0. \end{aligned}\right\} \tag{14}$$

La deuxième équation se déduit ici de la première, c'est pourquoi le domaine de solutions du système (14) est le plan $\mathcal{L}$ défini par l'équation

$$x - y + z = 0.$$

On voit aisément que ce plan coupe l'axe z en un seul point et, par conséquent, il n'est pas parallèle à l'axe z. Déterminons l'ensemble $\mathcal{K}_z$.

En posant dans le système (13) $x = 0$, $y = 0$, on obtient le système

$$\left.\begin{aligned} z + 1 \geqslant 0, \\ -z + 2 \geqslant 0, \end{aligned}\right\}$$

d'où il s'ensuit que

$$-1 \leqslant z \leqslant 2. \tag{15}$$

Ainsi, $\mathcal{K}$ est l'ensemble $\mathcal{K}_z + \mathcal{L}$ qui se compose des points de la forme

$$(0,\ 0,\ z) + (x,\ y,\ -x + y) = (x,\ y,\ z - x + y),$$

où x et y sont arbitraires et z vérifie les inégalités (15).

Nous allons formuler, pour conclure ce paragraphe, deux théorèmes généralisant les deux derniers théorèmes du § 5 à un cas tridimensionnel. La seule modification à apporter dans leurs énoncés consiste à remplacer le terme « plan » par celui « espace ».

● THÉORÈME. *Tout domaine polyèdre convexe (non vide) dans un espace peut être mis sous la forme de la somme suivante*

$$\langle A_1, A_2, \ldots, A_p \rangle + (B_1, B_2, \ldots, B_q).$$

● THÉORÈME. *Tout ensemble de la forme*

$$\langle A_1, A_2, \ldots, A_p \rangle + (B_1, B_2, \ldots, B_q)$$

dans un espace représente ou bien l'espace tout entier ou bien un certain domaine polyèdre convexe dans celui-ci.

Les démonstrations de ces théorèmes sont absolument identiques au cas bidimensionnel. Nous laissons au lecteur le soin de le faire.

§ 7. SYSTÈMES D'INÉGALITÉS LINÉAIRES À UN NOMBRE ARBITRAIRE D'INCONNUES

Jusqu'à présent nous nous sommes bornés à examiner les systèmes d'inégalités *à deux* ou *trois inconnues*. La raison en est que : premièrement, l'exploration de ces systèmes est facile et s'accorde bien avec le programme scolaire ; deuxièmement (et cela est d'ailleurs plus important), les solutions de ces systèmes admettent une interprétation géométrique fort concrète (les points dans un plan ou dans un espace). Mais dans les applications (en programmation linéaire par exemple), on rencontre souvent des systèmes d'inégalités à un nombre d'inconnues $n > 3$. Il serait donc déraisonnable de les passer sous silence. Nous examinerons ci-dessous, bien que brièvement d'ailleurs, certains systèmes à un nombre arbitraire d'inconnues ($n > 3$).

Pour une interprétation géométrique d'un système d'inégalités linéaires à n inconnues, il faut nous adresser à un espace à *n dimensions*.

Commençons par l'énoncé des notions les plus fondamentales.

Un *point d'un espace à n dimensions* est, par définition, donné par une collection réglementée de n nombres

$$x_1,\ x_2,\ \ldots,\ x_n$$

appelés *coordonnées* de ce point. Cet énoncé se fonde sur le fait que dans la géométrie analytique un point dans un plan se caractérise par deux nombres et dans un espace, par trois. Par la suite au lieu de dire « un point M a les coordonnées $x_1, x_2, \ldots, x_n$ » nous écrirons $M = (x_1, x_2, \ldots, x_n)$ ou $M(x_1, x_2, \ldots, x_n)$ tout court. Le point $(0, 0, \ldots, 0)$ est appelé *origine des coordonnées* ou *origine* tout simplement.

Définissons pour commencer ce qu'on entend par « segment » dans un espace à n dimensions. Selon la définition correspondante du § 1, un segment M_1M_2 dans un espace usuel se caractérise comme l'ensemble de tous les points de la forme

$$s_1M_1 + s_2M_2,$$

où s_1, s_2 sont deux nombres non négatifs quelconques dont la somme vaut 1. Ainsi, en passant de l'espace tridimensionnel à celui à n dimensions nous prenons cette caractéristique pour la définition d'un segment. Plus précisément, soient

$$M'(x_1', x_2', \ldots, x_n') \quad \text{et} \quad M''(x_1'', x_2'', \ldots, x_n'')$$

deux points arbitraires de l'espace à n dimensions. On appelle alors *segment M_1M_2 l'ensemble de tous les points de la forme*

$$s'M' + s''M'' = (s'x_1' + s''x_1'',\ s'x_2' + s''x_2'',\ \ldots,\ s'x_n' + s''x_n''), \quad (1)$$

où s', s'' sont deux nombres non négatifs quelconques dont la somme vaut 1. Pour $s' = 1$, $s'' = 0$ on obtient le point M', pour $s' = 0$, $s'' = 1$ le point M''. Ce sont les *extrémités* du segment $M'M''$. Les autres points (pour $s' > 0$, $s'' > 0$) sont dits *points intérieurs* d'un segment.

Dans ce qui suit nous aurons également besoin de notion d'hyperplan dans l'espace n-dimensionnel. C'est une généralisation de la notion de plan dans l'espace tridimensionnel usuel. Le préfixe « hyper » a ici un sens bien déterminé; le fait est que dans l'espace à n dimensions on distingue les « plans » différents: plans unidimensionnels (« lignes droites »), bidimensionnels, etc., et, enfin, à $(n - 1)$ dimensions; ces derniers s'appellent justement « hyperplans ».

● DÉFINITION. *On appelle hyperplan dans un espace à n dimensions l'ensemble des points* M $(x_1, x_2, \ldots, x_n)$ *dont les coordonnées satisfont à l'équation du premier degré*

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = 0, \tag{2}$$

où au moins l'un des nombres $a_1, a_2, \ldots, a_n$ (coefficients des inconnues) *n'est pas nul.* Pour $n = 3$ l'équation (2) prend la forme $a_1x_1 + a_2x_2 + a_3x_3 + b = 0$, ce qui n'est rien d'autre que l'équation d'un plan dans l'espace usuel (où les coordonnées sont désignées par x_1, x_2, x_3 et non pas x, y, z comme d'habitude).

Par rapport à l'hyperplan (2), tout l'espace à n dimensions se divise en deux parties: le domaine où a lieu l'inégalité

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \geqslant 0, \tag{3}$$

et celui dans lequel

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \leqslant 0. \tag{4}$$

Ces domaines sont appelés *demi-espaces*. De cette façon, chaque hyperplan partage tout l'espace en deux demi-espaces ayant cet hyperplan pour leur partie commune.

La notion de corps convexe se généralise également à l'espace à n dimensions. Un ensemble des points dans un espace à n dimensions est appelé *convexe* si tout en contenant ses deux points M' et M'' quelconques il contient également tout le segment $M'M''$.

Il est facile de montrer que *tout demi-espace est un ensemble convexe*. En effet, soient deux points M' $(x'_1, x'_2, \ldots, x'_n)$ et M'' $(x''_1, x''_2, \ldots, x''_n)$ appartenant à un demi-espace (3). Démontrons que tout point M du segment $M'M''$ appartient également à ce demi-espace.

Les coordonnées du point M sont de la forme (1) ou, ce qui revient au même, s'écrivent

$$\begin{aligned} x_1 &= sx'_1 + (1-s)\,x''_1, \\ x_2 &= sx'_2 + (1-s)\,x''_2, \qquad (0 \leqslant s \leqslant 1). \\ &\ldots\ldots\ldots\ldots \\ x_n &= sx'_n + (1-s)\,x''_n. \end{aligned}$$

En portant ces expressions dans le premier membre de (3), on obtient

$$\begin{aligned} a_1(sx'_1 + (1-s)\,x''_1) + a_2(sx'_2 + (1-s)\,x''_2) + \ldots \\ \ldots + a_n(sx'_n + (1-s)\,x''_n) + b = \\ = s(a_1x'_1 + a_2x'_2 + \ldots + a_nx'_n) + \\ + (1-s)(a_1x''_1 + a_2x''_2 + \ldots + a_nx''_n) + sb + (1-s)\,b \end{aligned}$$

(nous avons remplacé le nombre b par la somme $sb + (1-s)\,b$), mais cette dernière expression est égale à

$$s\,[a_1x'_1 + \ldots + a_nx'_n + b] + (1-s)\,[a_1x''_1 + \ldots + a_nx''_n + b].$$

Chacune des deux sommes mises entre crochets est non négative, car les points M' et M'' appartiennent tous les deux au demi-espace (3). Par conséquent, l'expression ci-dessus tout entière est également non négative (vu que $s \geqslant 0$ et $(1-s) \geqslant 0$). Ceci prouve que le point M appartient au demi-espace (3), c'est-à-dire que ce dernier est convexe.

Il est aisé maintenant d'introduire la terminologie géométrique correspondant à un système d'inégalités linéaires

à n inconnues. Soit un système

$$\left.\begin{array}{l} a_1x_1 + a_2x_2 + \ldots + a_nx_n + a \geqslant 0, \\ b_1x_1 + b_2x_2 + \ldots + b_nx_n + b \geqslant 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_1x_1 + c_2x_2 + \ldots + c_nx_n + c \geqslant 0. \end{array}\right\}$$

Chacune des inégalités ci-dessus définit un demi-espace et toutes les inégalités prises ensemble, un domaine $\mathcal{K}$ dans l'espace à n dimensions, ce domaine étant l'intersection d'un nombre fini de demi-espaces. Le domaine $\mathcal{K}$ étant formé de demi-espaces convexes l'est aussi.

Par analogie avec l'espace tridimensionnel, *on appelle domaine polyèdre convexe tout domaine dans l'espace à* n *dimensions représentant l'intersection d'un nombre fini de demi-espaces et polyèdre convexe tout court si cette intersection est un ensemble limité.* Il faut comprendre le terme « ensemble limité » dans ce sens que les coordonnées de tous les points du domaine considéré ne dépassent pas en valeur absolue une certaine constante c:

$$|x_1| \leqslant c, \ \ldots, \ |x_n| \leqslant c \text{ pour tous les points du domaine considéré.}$$

Ainsi, *l'ensemble des points d'un espace à* n *dimensions les coordonnées desquels satisfont au système* (5) *est un domaine polyèdre convexe* $\mathcal{K}$ *qui s'obtient comme l'intersection de tous les demi-espaces correspondant aux inégalités du système donné.*

Il est à souligner une fois de plus que si ce domaine est limité il est appelé polyèdre convexe.

Les méthodes de description du domaine $\mathcal{K}$ examinées pour les systèmes à deux inconnues (§ 5) et à trois inconnues (§ 6) sont étendues, avec des modifications correspondantes, aux systèmes à n inconnues. Comme leur exposé détaillé prendrait ici beaucoup de place, nous ne nous attarderons pas là-dessus. Il est de plus à souligner que pour un grand nombre d'inconnues, ces méthodes deviennent peu efficaces et entraînent des calculs fort encombrants.

Il est à remarquer que les théorèmes généraux sur la structure des ensembles polyèdres convexes dans l'espace tridimensionnel restent entièrement en vigueur dans l'espace à n dimensions bien que leurs démonstrations deviennent plus compliquées. Nous ne nous bornerons qu'à leur énoncé et aux explications nécessaires.

● THÉORÈME 1. *L'enveloppe convexe de tout système fini de points* $A_1, A_2, \ldots, A_p$ *est un polyèdre convexe.*

Pour plus de clarté, il convient de souligner que par ceci on établit le rapport entre les deux types des ensembles définis de manière tout à fait différente: entre l'enveloppe convexe d'un système de points $A_1, A_2, \ldots, A_p$ qui se note $\langle A_1, A_2, \ldots, A_p\rangle$ et se définit comme l'ensemble de tous les points de la forme

$$s_1A_1 + s_2A_2 + \ldots + s_pA_p,$$

où $s_1, s_2, \ldots, s_p$ sont des nombres non négatifs quelconques dont la somme vaut 1, et les polyèdres convexes, c'est-à-dire des domaines limités qui s'obtiennent comme l'intersection d'un nombre fini de demi-espaces.

Dans l'espace bidimensionnel et dans celui tridimensionnel le bien-fondé du théorème 1 est évident (à en juger ne serait-ce que par le sens concret de l'enveloppe convexe) tandis que dans l'espace à n dimensions il n'est pas tellement apparent et exige nécessairement une démonstration.

● THÉORÈME 1' (réciproque de 1). *Tout polyèdre convexe coïncide avec l'enveloppe convexe d'un système fini de points.*

En réalité, on peut affirmer même davantage: *un polyèdre convexe se confond avec l'enveloppe convexe de ses sommets.* La définition d'un sommet dans l'espace à n dimensions est absolument la même que dans l'espace bidimensionnel (le sommet est un point du polyèdre tel qui n'est pas intérieur pour aucun segment entièrement compris dans ce polyèdre). On peut bien montrer que le nombre de sommets est toujours fini.

● THÉORÈME 2. *Tout ensemble de la forme* $(B_1, B_2, \ldots, B_q)$ *ou bien se confond avec l'espace tout entier, ou bien représente un cône polyèdre convexe dont le sommet coïncide avec l'origine des coordonnées.*

Rappelons que le symbole $(B_1, B_2, \ldots, B_q)$ désigne l'ensemble de tous les points dont la forme est

$$t_1B_1 + t_2B_2 + \ldots + t_qB_q,$$

où $t_1, t_2, \ldots, t_q$ sont des nombres non négatifs quelconques. Un cône polyèdre convexe s'obtient comme l'intersection d'un nombre fini de demi-espaces dont les hyperplans limités ont un point commun

(sommet du cône). Le bien-fondé du théorème 2 dans un espace tridimensionnel a été démontré au § 4 (théorème 1).

● THÉORÈME 2′. *Tout cône polyèdre convexe ayant l'origine des coordonnées pour sommet peut se représenter sous la forme* $(B_1, B_2, \ldots, B_q)$.

Le bien-fondé du théorème pour le cas tridimensionnel a été démontré plus haut (théorème 2 du § 4).

● THÉORÈME 3. *Tout domaine polyèdre convexe peut se représenter sous la forme d'une somme*

$$\langle A_1, A_2, \ldots, A_p\rangle + (B_1, B_2, \ldots, B_q).$$

● THÉORÈME 3′. *Toute somme de la forme ci-dessus représente ou bien l'espace tout entier ou bien un domaine polyèdre convexe dans celui-ci.*

§ 8. SYSTÈMES INCOMPATIBLES

Jusqu'à présent nous nous sommes intéressés tout particulièrement aux systèmes d'inégalités ayant *ne serait-ce qu'une solution* (systèmes compatibles). Les domaines correspondants (dans un plan ou dans un espace) représentaient des ensembles non vides des points. Quant aux systèmes incompatibles leur étude semble, à première vue, inutile; de plus, il semble peu probable qu'on puisse lier ces systèmes à une théorie quelque peu intéressante et de valeur. Mais c'est toujours « à première vue ». En réalité voilà en sont les choses: les propriétés des systèmes incompatibles présentent de l'intérêt non seulement par elles-mêmes mais elles aident à comprendre toute une série de faits importants. Ainsi, par exemple, le théorème fondamental de la programmation linéaire (théorème de dualité, voir § 10) se déduit finalement à partir de certaines propriétés de tels systèmes.

Considérons un système quelconque d'inégalités linéaires. Admettons, pour plus de commodité, que le nombre d'inconnues soit égal à 3, bien que tous nos raisonnements restent évidemment valables pour des systèmes à un nombre d'inconnues quelconque.

Soit un système

$$\left.\begin{aligned} a_1x + b_1y + c_1z + d_1 &\geqslant 0, \\ a_2x + b_2y + c_2z + d_2 &\geqslant 0, \\ \dots\dots\dots\dots\dots\dots \\ a_mx + b_my + c_mz + d_m &\geqslant 0. \end{aligned}\right\} \quad (1)$$

Multiplions les deux membres de la première inégalité (1) par un nombre non négatif k_1 quelconque, les deux membres de la deuxième par k_2, etc., pour additionner ensuite les inégalités ainsi obtenues. On obtient donc l'inégalité

$$\begin{aligned} (k_1a_1 + k_2a_2 + \dots + k_ma_m)\, x + \\ + (k_1b_1 + k_2b_2 + \dots + k_mb_m)\, y + \\ + (k_1c_1 + k_2c_2 + \dots + k_mc_m)\, z + \\ + k_1d_1 + k_2d_2 + \dots + k_md_m \geqslant 0, \quad (2) \end{aligned}$$

que nous appellerons *combinaison linéaire des inégalités* (1).

Il peut s'avérer qu'une certaine combinaison linéaire des inégalités (1) représente une inégalité de la forme

$$0 \cdot x + 0 \cdot y + 0 \cdot z + d \geqslant 0, \quad (3)$$

où d est un nombre négatif dans le sens strict de ce terme (après la division par $| d |$ on obtient alors l'inégalité $-1 \geqslant 0$). Il est clair qu'aucune collection de valeurs des inconnues ne vérifiera une telle inégalité, c'est pourquoi dans le cas examiné le système (1) est incompatible (n'a pas de solutions). Il est remarquable que la réciproque a également lieu, à savoir: si le système (1) est incompatible, une certaine combinaison linéaire de ses inégalités a alors la forme (3).

Nous allons démontrer cette proposition sous sa forme générale (c'est-à-dire pour des systèmes à un nombre d'inconnues quelconque). Nous commençons par introduire la *définition* suivante. Convenons d'appeler *contradictoire* l'inégalité

$$ax + by + cz + d \geqslant 0$$

si celle-ci n'est pas vérifiée pour aucune collection de valeurs des inconnues. Toute inégalité contradictoire a évidemment la forme (3), où $d < 0$. Ainsi, la proposition que nous avons à démontrer peut s'énoncer comme suit :

Théorème sur les systèmes incompatibles d'inégalités. Si un système d'inégalités linéaires est incompatible, alors une certaine combinaison linéaire de celles-ci est une inégalité contradictoire.

La *démonstration* se fera par récurrence sur n (nombre d'inconnues du système considéré).

Pour $n = 1$ le système considéré aura pour expression

$$\left.\begin{array}{r} a_1x + b_1 \geqslant 0, \\ a_2x + b_2 \geqslant 0, \\ \cdots\cdots\cdots \\ a_mx + b_m \geqslant 0. \end{array}\right\} \qquad (4)$$

On peut considérer que tous les coefficients $a_1, a_2, \ldots, a_m$ ne sont pas nuls. En effet, si $a_1 = 0$, par exemple, la première inégalité prend la forme $0 \cdot x + b_1 \geqslant 0$; si b_1 est non négatif, cette inégalité est à rejeter et si b_1 est négatif, alors la première inégalité est contradictoire et il n'y a rien à démontrer.

Considérons donc qu'aucun des nombres $a_1, a_2, \ldots, a_m$ n'est nul. Il est aisé de voir que parmi ces nombres il y en a aussi bien positifs que négatifs ; en effet, si tous ces nombres avaient le même signe, s'ils étaient positifs, par exemple, le système (4) se réduirait à la forme

$$\left.\begin{array}{l} x \geqslant -\dfrac{b_1}{a_1}, \\ x \geqslant -\dfrac{b_2}{a_2}, \\ \cdots\cdots\cdots \\ x \geqslant -\dfrac{b_m}{a_m}, \end{array}\right\}$$

et, par conséquent, il serait compatible.

Supposons, pour fixer les idées, que les premiers k des nombres $a_1, a_2, \ldots, a_m$ soient positifs et les autres $m - k$, négatifs. Le système (4) sera alors équivalent à celui

$$\left.\begin{array}{|c|}\hline x \geqslant -\dfrac{b_1}{a_1}; \\ \ldots\ldots\ldots\ldots \\ x \geqslant -\dfrac{b_k}{a_k}, \\ \hline \end{array}\right. \\ \left.\begin{array}{|c|}\hline x \leqslant -\dfrac{b_{k+1}}{a_{k+1}}, \\ \ldots\ldots\ldots\ldots \\ x \leqslant -\dfrac{b_m}{a_m}. \\ \hline \end{array}\right\} \qquad (5)$$

Choisissons parmi les nombres $-\frac{b_1}{a_1}, \ldots, -\frac{b_k}{a_k}$ celui qui est le plus grand. Admettons que ce soit $-\frac{b_1}{a_1}$. Dans ce cas, les k premières inégalités du système (5) peuvent être remplacées par la première. De façon analogue, choisissons parmi les nombres $-\frac{b_{k+1}}{a_{k+1}}, \ldots, -\frac{b_m}{a_m}$ celui le plus petit. Les autres $m - k$ inégalités du système (5) ne peuvent être remplacées que par la dernière inégalité. De cette façon, le système (4) est équivalent au système suivant de deux inégalités:

$$\left.\begin{array}{l} x \geqslant -\dfrac{b_1}{a_1}, \\ x \leqslant -\dfrac{b_m}{a_m}, \end{array}\right\}$$

et son incompatibilité signifie que

$$-\frac{b_1}{a_1} > -\frac{b_m}{a_m}, \qquad (6)$$

De (6) s'ensuit que

$$b_m a_1 - b_1 a_m < 0 \qquad (7)$$

(il faut tenir compte que $a_1 > 0$ et $a_m < 0$). Si l'on multiplie maintenant la première des inégalités (4) par le nombre positif a_m et la dernière par le nombre positif a_1 pour additionner ensuite les inégalités obtenues, on trouvera alors l'inégalité

$$0 \cdot x + (b_m a_1 - b_1 a_m) \geqslant 0$$

qui est contradictoire en vertu de (7). Ainsi, le théorème est vrai pour un système à une inconnue.

Supposons maintenant que ce théorème soit vrai pour les systèmes d'inégalités à $n - 1$ inconnues et démontrons sous cette hypothèse sa validité pour un système à n inconnues.

Soit un système incompatible d'inégalités linéaires dont les inconnues sont $x_1, x_2, \ldots, x_n$. Considérons n'importe laquelle de ces inégalités. Celle-ci a la forme

$$a_1 x_1 + \ldots + a_{n-1} x_{n-1} + a_n x_n + b \geqslant 0$$

ou, après le report de $a_n x_n$ dans le deuxième membre,

$$a_1 x_1 + \ldots + a_{n-1} x_{n-1} + b \geqslant -a_n x_n.$$

Si $a_n = 0$ on laisse cette inégalité inchangée. Si $a_n < 0$ on divise les deux membres de celle-ci par un nombre positif a_n; on obtient l'inégalité de la forme

$$a'_1 x_1 + \ldots + a'_{n-1} x_{n-1} + b' \geqslant x_n.$$

Pour $a_n > 0$, divisons les deux membres de cette inégalité par a_n. On obtient alors l'inégalité

$$-(a'_1 x_1 + \ldots + a'_{n-1} x_{n-1} + b') \geqslant -x_n.$$

Ainsi, en multipliant chacune des inégalités du système initial par un nombre positif convenablement choisi, on

obtient un système équivalent de la forme

$$\left.\begin{array}{l} \boxed{\begin{array}{c} P_1 \geqslant x_n, \\ P_2 \geqslant x_n, \\ \cdots\cdots \\ P_p \geqslant x_n; \end{array}} \\ \\ \boxed{\begin{array}{c} -Q_1 \geqslant -x_n, \\ -Q_2 \geqslant -x_n, \\ \cdots\cdots \\ -Q_q \geqslant -x_n; \end{array}} \\ \\ \boxed{\begin{array}{c} R_1 \geqslant 0, \\ R_2 \geqslant 0, \\ \cdots\cdots \\ R_r \geqslant 0, \end{array}} \end{array}\right\} \quad (8)$$

où $P_1, \ldots, P_p, Q_1, \ldots, Q_q, R_1, \ldots, R_r$ sont les expressions de la forme $a_1x_1 + \ldots + a_{n-1}x_{n-1} + b$ (ne contenant pas x_n).

Par hypothèse le système initial est incompatible. D'où il s'ensuit immédiatement que le système (8) l'est aussi. Il en résulte l'incompatibilité du système

$$\left.\begin{array}{l} \boxed{\begin{array}{c} P_1 \geqslant Q_1, \\ \cdots\cdots \\ P_p \geqslant Q_q; \end{array}} \\ \\ \boxed{\begin{array}{c} R_1 \geqslant 0, \\ \cdots\cdots \\ R_r \geqslant 0. \end{array}} \end{array}\right\} \quad (9)$$

aux inconnues $x_1, \ldots, x_{n-1}$ (la partie supérieure encadrée de ce système se compose des inégalités de la forme $P_\alpha \geqslant$

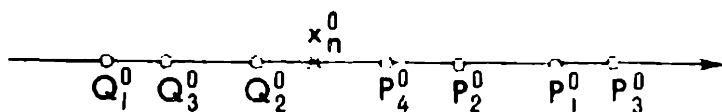


Fig. 46

$\geqslant Q_\beta$, où α est un nombre quelconque des nombres 1, 2, ..., p, β étant un nombre arbitraire des nombres 1, 2, ..., q). En effet, si le système (9) était compatible cela signifierait que pour certaines valeurs des inconnues

$$x_1 = x_1^0, \ldots, x_{n-1} = x_{n-1}^0$$

ont lieu les inégalités (numériques) suivantes

$$\boxed{\begin{array}{c} P_1^0 \geqslant Q_1^0, \\ \cdot\ \cdot\ \cdot\ \cdot \\ P_p^0 \geqslant Q_q^0; \end{array}}$$

$$\boxed{\begin{array}{c} R_1^0 \geqslant 0, \\ \cdot\ \cdot\ \cdot \\ R_r^0 \geqslant 0, \end{array}}$$

où P_α^0 représente la valeur de P_α pour $x_1 = x_1^0, \ldots, x_{n-1} = x_{n-1}^0$ (Q_β^0 et R_γ^0 ont le sens identique). Ainsi, chacun des nombres $Q_1^0, \ldots, Q_q^0$ ne dépasse pas n'importe lequel des nombres $P_1^0, \ldots, P_p^0$. Mais dans ce cas il existe nécessairement un nombre x_n^0 compris entre tous les nombres $Q_1^0, \ldots, Q_q^0$ et $P_1^0, \ldots, P_p^0$:

$$P_1^0 \geqslant x_n^0 \geqslant Q_1,$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$
$$P_p^0 \geqslant x_n^0 \geqslant Q_q^0$$

(fig. 46). Les inégalités qui en découlent

$$\boxed{\begin{array}{c} P_1^0 \geqslant x_n^0, \\ \cdots\cdots \\ P_p^0 \geqslant x_n^0 \end{array}}$$

$$\boxed{\begin{array}{c} -Q_1^0 \geqslant -x_n^0, \\ \cdots\cdots \\ -Q_q^0 \geqslant -x_n^0 \end{array}}$$

et celles

$$\boxed{\begin{array}{c} R_1^0 \geqslant 0, \\ \cdots\cdots \\ R_r^0 \geqslant 0 \end{array}}$$

montrent que la collection des valeurs des inconnues

$$x_1 = x_1^0, \ldots, x_{n-1} = x_{n-1}^0, \quad x_n = x_n^0$$

représente la solution du système (8), ce qui contredit notre hypothèse (le système (8) est incompatible).

Ainsi, le système (9) est incompatible. Mais vu que le nombre d'inconnues de ce système est $n - 1$, on peut donc lui appliquer la méthode de récurrence. Cela signifie qu'une certaine combinaison linéaire des inégalités (9) représente une inégalité contradictoire. Il est aisé de voir que chacune des inégalités (9) est une combinaison linéaire des inégalités (8) : en effet, si l'on additionne tout simplement les inégalités $P_\alpha \geqslant x_n$ et $-Q_\beta \geqslant -x_n$ du système (8), on obtient $P_\alpha - Q_\beta \geqslant 0$ ou $P_\alpha \geqslant Q_\beta$, c'est-à-dire une des inégalités (9). Par conséquent, une certaine combinaison linéaire des inégalités (8) représente elle aussi une inégalité contradictoire. D'où il s'ensuit qu'une certaine combinaison liné-

aire des inégalités du système initial est également une inégalité contradictoire. Le théorème est démontré.

Le théorème sur l'incompatibilité des systèmes d'inégalités linéaires n'est qu'une des manifestations de l'analogie profonde qui existe entre les propriétés des systèmes d'inégalités linéaires et celles des *systèmes d'équations linéaires*. Essayons de remplacer dans l'énoncé du théorème le mot « inégalité » par « équation ». On obtient donc la proposition suivante:

Si un système d'équations linéaires est incompatible, alors une certaine combinaison linéaire de celles-ci est une équation contradictoire.

Cette proposition s'avère également juste. Enoncée sous une forme quelque peu différente, elle est dite *théorème de Kronecker-Kapelli*. On le démontre en *algèbre linéaire* (partie des mathématiques où l'on étudie les opérations linéaires, c'est-à-dire celles semblables à l'addition des points et à la multiplication d'un point par un nombre quelconque dans l'espace à n dimensions). Pour une assimilation correcte de ce qu'on a dit plus haut, nous devons préciser quelque peu la notion de combinaison linéaire. Une combinaison linéaire d'équations se forme de la même manière que celle d'inégalités à cette différence seulement que l'on peut multiplier les équations considérées par *tout nombre* et non seulement par un nombre non négatif. Tout comme pour les inégalités, une équation est dite contradictoire si elle n'a pas de solutions. Il est facile de montrer qu'une équation contradictoire est nécessairement de la forme

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n + b = 0,$$

b étant un nombre quelconque non nul (après la division des deux membres de cette équation par b on obtient l'« équation » $1 = 0$).

Un cas particulier du théorème sur les systèmes incompatibles d'inégalités est d'une importance particulière,

celui où le système considéré contient des inégalités

$$x_1 \geqslant 0,\ x_2 \geqslant 0,\ \ldots,\ x_n \geqslant 0. \qquad (10)$$

Après avoir désigné par (S) la partie restante du système considéré, on peut dire que le problème consiste à rechercher toutes les solutions *non négatives* (c'est-à-dire vérifiant les conditions (10)) du système (S). Si ce problème n'a pas de solutions, alors en vertu du théorème qui vient d'être démontré une certaine combinaison linéaire des inégalités du système (S)

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + a \geqslant 0 \qquad (11)$$

avec une combinaison linéaire quelconque des inégalités (10)

$$k_1x_1 + k_2x_2 + \ldots + k_nx_n \geqslant 0$$

($k_1, k_2, \ldots, k_n$ étant non négatifs)

donnent l'inégalité contradictoire suivante

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n + c \geqslant 0,$$

où c est un nombre négatif quelconque. Par conséquent,

$$a_1 = -k_1 \leqslant 0,\ a_2 = -k_2 \leqslant 0,\ \ldots,\ a_n = -k_n \leqslant 0,\quad a < 0.$$

Enonçons le résultat obtenu sous la forme d'une proposition isolée.

Corollaire du théorème sur les systèmes incompatibles. Si un système d'inégalités n'a pas de solutions non négatives, une certaine combinaison linéaire de ces inégalités représente une inégalité de la forme (11), *où tous les coefficients sont* $a_1, a_2, \ldots, a_n \leqslant 0$, *et le terme libre* $a < 0$.

Un autre corollaire fort important de ce théorème, c'est le lien que l'on peut établir entre le système d'inégalités donné et un autre système qui, en plus des inégalités, contient des *équations*. Eclaircissons ce lien à l'aide du système (1) (à trois inconnues x, y, z).

Si le système (1) est incompatible, une certaine combinaison linéaire (2) de ses inégalités est de la forme (3), où $d < 0$. Cela montre qu'il existe des nombres non négatifs $k_1, k_2, \ldots, k_m$ tels que

$$k_1a_1 + k_2a_2 + \ldots + k_ma_m = 0,$$
$$k_1b_1 + k_2b_2 + \ldots + k_mb_m = 0,$$
$$k_1c_1 + k_2c_2 + \ldots + k_mc_m = 0,$$
$$k_1d_1 + k_2d_2 + \ldots + k_md_m < 0.$$

En d'autres termes, cela veut dire que le système mixte

$$\left.\begin{array}{l} a_1y_1 + a_2y_2 + \ldots + a_my_m = 0, \\ b_1y_1 + b_2y_2 + \ldots + b_my_m = 0, \\ c_1y_1 + c_2y_2 + \ldots + c_my_m = 0, \\ d_1y_1 + d_2y_2 + \ldots + d_my_m = -1, \\ y_1 \geqslant 0, \\ y_2 \geqslant 0, \\ \ldots\ldots\ldots\ldots \\ y_m \geqslant 0, \end{array}\right\} \qquad (1')$$

qui se compose des équations et des inégalités, a pour solution $y_1 = \lambda k_1$, $y_2 = \lambda k_2, \ldots, y_n = \lambda k_n$, c'est-à-dire qu'il est compatible *).

Ainsi, *si le système* (1) *est incompatible, celui* (1') *est, par contre, compatible. La réciproque est certainement vraie* (*si le système* (1') *est compatible,* (1) *ne l'est pas*), car la compatibilité du système (1') signifie qu'une certaine combinaison linéaire des inégalités (1) est une inégalité contradictoire $-1 \geqslant 0$.

§ 9. CÔNES POLYÈDRES CONVEXES RÉCIPROQUEMENT DUAUX

Examinons maintenant plus en détail les cônes polyèdres convexes mentionnés au § 4.

Chaque cône polyèdre convexe de sommet à l'origine des coordonnées représente dans un espace tridimensionnel

*) Le nombre λ est choisi de façon que l'on ait $d_1(\lambda k_1) + d_2(\lambda k_2) + \ldots + d_m(\lambda k_m) = -1$, c'est-à-dire

$$\lambda = -\frac{1}{d_1k_1 + d_2k_2 + \ldots + d_mk_m}.$$

comme nous l'avons montré plus haut, le domaine de solutions d'un système d'inégalités linéaires *homogènes* à trois inconnues:

$$\left.\begin{aligned} a_1x + b_1y + c_1z \geqslant 0, \\ a_2x + b_2y + c_2z \geqslant 0, \\ a_mx + b_my + c_mz \geqslant 0. \end{aligned}\right\} \qquad (1)$$

Soit, en plus de ce système, une inégalité

$$ax + by + cz \geqslant 0. \qquad (2)$$

On dit que l'inégalité (2) *est une conséquence du système* (1) *si toute collection de valeurs des inconnues* x, y, z, *vérifiant le système* (1), *satisfait également à l'inégalité* (2).

Certes, toute inégalité représentant une combinaison linéaire des inégalités (1) est une conséquence du système (1). La question se pose naturellement de savoir: la réciproque est-elle juste? Il s'avère que oui.

● THÉORÈME 1. *L'inégalité homogène* (2), *conséquence du système homogène* (1), *peut se représenter comme une combinaison linéaire des inégalités* (1).

● DÉMONSTRATION. Désignons, pour plus de commodité, par P_1, P_2, . . ., P_m les membres gauches des première, deuxième, . . ., m-ième inégalités respectivement du système (1) et par P le membre gauche de l'inégalité (2). Ainsi, nous avons un système

$$\left.\begin{aligned} P_1 \geqslant 0, \\ P_2 \geqslant 0, \\ \cdots\cdots \\ P_m \geqslant 0, \end{aligned}\right\} \qquad (1)$$

ainsi qu'une inégalité

$$P \geqslant 0, \qquad (2)$$

conséquence du système (1). Il est à démontrer que cette inégalité peut s'écrire comme une combinaison linéaire des inégalités (1).

Vu que l'inégalité $P \geqslant 0$ est une conséquence du système (1), l'équation $P = -1$ est, par contre, incompatible avec celui-ci ; en d'autres termes, le système mixte

$$\left.\begin{aligned} P_1 &\geqslant 0, \\ P_2 &\geqslant 0, \\ &\cdots\cdots \\ P_m &\geqslant 0, \\ P &= -1 \end{aligned}\right\} \qquad (3)$$

est incompatible. Essayons donc d'appliquer à ce dernier le théorème sur les systèmes incompatibles. On ne peut pas le faire de façon directe puisque ce théorème se rapporte aux systèmes qui ne contiennent que des *inégalités* tandis que le système (3) contient de plus une *équation* $P = -1$. Mais cette équation-là est elle-même équivalente au système

$$\left.\begin{aligned} P &\geqslant -1, \\ P &\leqslant -1 \end{aligned}\right\} \quad \text{ou, ce qui revient au même,} \quad \left.\begin{aligned} P + 1 &\geqslant 0, \\ -P - 1 &\geqslant 0, \end{aligned}\right\}$$

d'où il s'ensuit que le système (3) est équivalent à

$$\left.\begin{aligned} P_1 &\geqslant 0, \\ P_2 &\geqslant 0, \\ &\cdots\cdots \\ P_m &\geqslant 0, \\ P + 1 &\geqslant 0, \\ -P - 1 &\geqslant 0, \end{aligned}\right\} \qquad (4)$$

qui s'avère incompatible tout comme (3).

En vertu du théorème sur les systèmes incompatibles, la combinaison linéaire des inégalités (4) est une inégalité

contradictoire. En d'autres termes, il existe des nombres non négatifs $k_1, k_2, \ldots, k_m, k_{m+1}, k_{m+2}$ tels que l'inégalité

$$k_1P_1 + k_2P_2 + \ldots + k_mP_m + k_{m+1}(P+1) + \\ + k_{m+2}(-P-1) \geqslant 0$$

prendra (après la réduction des termes semblables) la forme

$$0 \cdot x + 0 \cdot y + 0 \cdot z + d \geqslant 0,$$

d étant un nombre négatif. Par conséquent,

$$k_1P_1 + k_2P_2 + \ldots + k_mP_m + (k_{m+1} - k_{m+2})P = \\ = 0 \cdot x + 0 \cdot y + 0 \cdot z,$$

et le nombre $k_{m+1} - k_{m+2} = d$ est négatif. D'où l'on obtient

$$P = \frac{k_1}{k_{m+2} - k_{m+1}} P_1 + \ldots + \frac{k_m}{k_{m+2} - k_{m+1}} P_m,$$

les facteurs de $P_1, \ldots, P_m$ étant non négatifs. Il est aisé de voir que cela signifie que l'inégalité (2) est une combinaison linéaire des inégalités (1), ce qu'il fallait démontrer.

Le théorème démontré présente par lui-même un intérêt mais son *interprétation géométrique* s'avère encore plus intéressante. Pour la faire connaître au lecteur, il nous faut aborder certains problèmes de la géométrie analytique aussi élémentaires d'ailleurs que quelques faits de celle-ci auxquels nous avons eu recours plus haut.

Soient

$$A(x_A, y_A, z_A), B(x_B, y_B, z_B)$$

deux points dans l'espace ne coïncidant pas avec l'origine des coordonnées O.

Appliquons au triangle OAB (fig. 47) le « théorème des cosinus », à savoir: le carré de tout côté d'un triangle est égal à la somme des carrés des deux autres côtés moins le produit double de ces mêmes côtés par le cosinus de l'angle

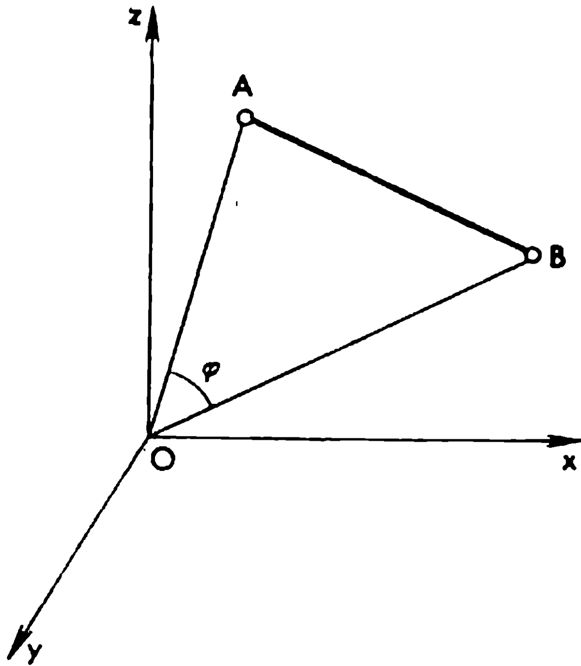


Fig. 47

compris entre eux. Dans le cas considéré on a

$$AB^2 = OA^2 + OB^2 - 2OA \cdot OB \cdot \cos\varphi, \qquad (5)$$

où φ est l'angle compris entre les segments OA et OB. Mais vu que

$$OA^2 = x_A^2 + y_A^2 + z_A^2,$$
$$OB^2 = x_B^2 + y_B^2 + z_B^2,$$
$$AB^2 = (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2,$$

après la réduction des termes semblables, on obtient de (5)

$$-2\,(x_A x_B + y_A y_B + z_A z_B) = -2\,OA \cdot OB \cdot \cos\varphi. \qquad (6)$$

L'angle φ sera non aigu si, et seulement si, $\cos\varphi \leqslant 0$. De (6) il s'ensuit également que

L'angle compris entre les segments OA et OB sera non aigu si, et seulement si,

$$x_A x_B + y_A y_B + z_A z_B \leqslant 0.$$

Convenons de désigner par la suite par (A, B) l'expression qui figure dans le membre gauche de cette inégalité:

$$(A, B) = x_A x_B + y_A y_B + z_A z_B.$$

Ceci dit, on peut formuler l'affirmation ci-dessus comme suit:

L'angle compris entre les segments OA et OB sera non aigu si, et seulement si,

$$(A, B) \leqslant 0.$$

Pour terminer notre petite digression dans le domaine de la géométrie analytique, arrêtons-nous brièvement sur une propriété de la grandeur (A, B), à savoir:

$$(k_1 A_1 + k_2 A_2, B) = k_1(A_1, B) + k_2 (A_2, B) \qquad (7)$$

quels que soient k_1 et k_2. La démonstration est presque évidente: puisque le point $k_1 A_1 + k_2 A_2$ a pour ses coordonnées

$$k_1 x_{A_1} + k_2 x_{A_2},\ k_1 y_{A_1} + k_2 y_{A_2},\ k_1 z_{A_1} + k_2 z_{A_2},$$

il vient

$$\begin{aligned}(k_1 A_1 + k_2 A_2, B) &= \\ &= (k_1 x_{A_1} + k_2 x_{A_2}) x_B + (k_1 y_{A_1} + k_2 y_{A_2}) y_B + \\ &+ (k_1 z_{A_1} + k_2 z_{A_2}) z_B = k_1 (x_{A_1} x_B + y_{A_1} y_B + \\ &+ z_{A_1} z_B) + k_2 (x_{A_2} x_B + y_{A_2} y_B + z_{A_2} z_B) = \\ &= k_1 (A_1, B) + k_2 (A_2, B).\end{aligned}$$

Passons enfin à l'objet principal de ce paragraphe: aux cônes polyèdres convexes dans un espace. Par définition (§ 4), on appelle cône polyèdre convexe de sommet au point S l'intersection d'un nombre fini de demi-espaces dont les plans limites passent par S. L'exemple le plus typique d'un cône polyèdre convexe est, comme on le sait, une pyramide convexe infinie. Dans ce qui suit nous admettrons que le sommet S coïncide avec l'origine des coordonnées O.

Supposons qu'un point B, différant de l'origine des coordonnées, possède la propriété suivante: *le segment OB forme avec n'importe quel segment OA un angle non aigu, A étant un point arbitraire du cône considéré* $\mathcal{K}$. On peut toujours trouver un tel point B: il suffit pour cela de mener par le sommet du cône un plan π de façon que le cône se trouve entièrement compris dans l'un des deux demi-espaces ainsi formés (fig. 48). La perpendiculaire sur le plan π prolongée dans l'*autre* demi-espace se compose entièrement des points B cherchés.

Considérons l'ensemble de *tous* les points B possédant la propriété susmentionnée. Complétons cet ensemble encore par un point, origine des coordonnées, et désignons-le par $\mathcal{K}^*$. Démontrons tout d'abord la validité du lemme suivant:

● LEMME. $\mathcal{K}^*$ *est également un cône polyèdre convexe* (voir fig. 49).

● DÉMONSTRATION. En vertu du théorème 2 (§ 4), tout cône polyèdre convexe $\mathcal{K}$ représente un ensemble de la forme $(A_1, A_2, \ldots, A_m)$ (les désignations adoptées dans le théorème en question sont quelque peu différentes: $B_1, B_2, \ldots, B_q$ au lieu de $A_1, A_2, \ldots, A_m$ dans le cas considéré). Cela signifie que tout point $A \in \mathcal{K}$ peut être mis sous la forme

$$A = t_1A_1 + t_2A_2 + \ldots + t_mA_m,$$

$t_1, t_2, \ldots, t_m$ étant non négatifs. Si le point B appartient à l'ensemble $\mathcal{K}^*$, l'angle compris entre le segment OB et n'importe lequel des segments OA, $A \in \mathcal{K}$, est non aigu, c'est-à-dire

$$(A, B) \leqslant 0 \text{ pour tous les } A \in \mathcal{K}.$$

Comme

$$(A, B) = t_1 (A_1, B) + t_2 (A_2, B) + \ldots + t_m (A_m, B)$$

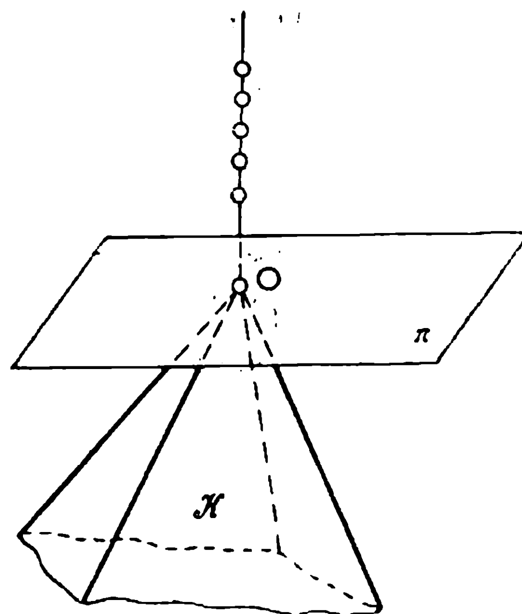


Fig. 48

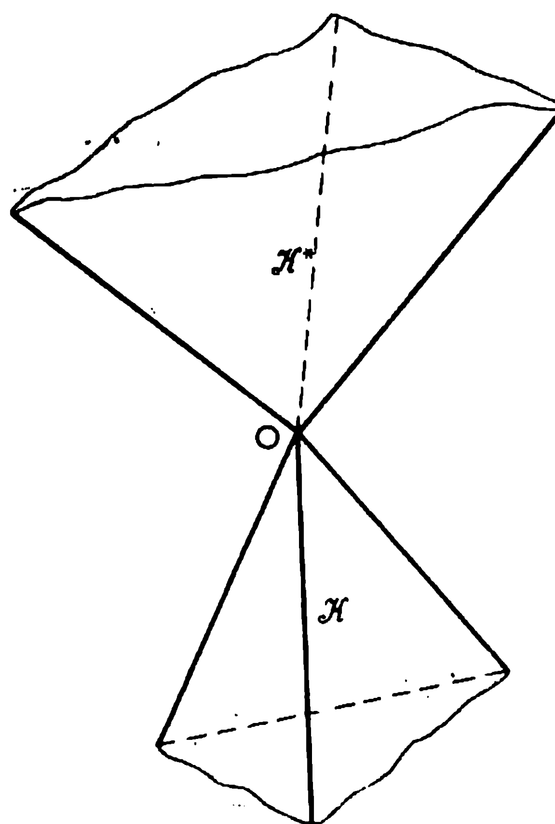


Fig. 49

(cf. formule (7)), on obtient donc

$$t_1 (A_1, B) + t_2 (A_2, B) + \ldots + t_m (A_m, B) \leqslant 0 \quad (8)$$

quels que soient les nombres $t_1, t_2, \ldots, t_m$ non négatifs. En particulier,

$$(A_1, B) \leqslant 0, \ (A_2, B) \leqslant 0, \ \ldots, \ (A_m, B) \leqslant 0. \quad (9)$$

Inversement, si les inégalités (9) sont vraies, l'inégalité (8) l'est aussi (pour tous les $t_1, t_2, \ldots, t_m$ non négatifs), c'est-à-dire $B \in \mathscr{K}^*$. Ainsi, *l'observation des inégalités* (9) *est nécessaire et suffisante pour que le point B appartienne à $\mathscr{K}^*$.*

Désignons par a_i, b_i, c_i $(i = 1, 2, \ldots, m)$ les coordonnées du point A_i et par x, y, z celles du point B. Les conditions (9) pourront alors s'écrire

$$\left.\begin{array}{r} a_1x + b_1y + c_1z \leqslant 0, \\ a_2x + b_2y + c_2z \leqslant 0, \\ \ldots\ldots\ldots\ldots\ldots \\ a_mx + b_my + c_mz \leqslant 0. \end{array}\right\} \quad (10)$$

Pour que le point B appartienne à l'ensemble $\mathscr{K}^*$, il est nécessaire et suffisant que ses coordonnées x, y, z satisfassent à toutes les inégalités (10). En d'autres termes:

$\mathscr{K}^*$ est le domaine de solutions du système (10).

Vu que le système (10) est homogène, le domaine de ses solutions est un cône polyèdre convexe dans un espace. Par conséquent, $\mathscr{K}^*$ est un cône polyèdre convexe, ce qu'il fallait démontrer.

Ainsi, *à chaque cône polyèdre convexe $\mathscr{K}$ nous pouvons maintenant faire correspondre au autre cône polyèdre convexe $\mathscr{K}^*$ qui se compose des points B tels pour lesquels le segment OB forme avec n'importe lequel des segments OA $(A \in \mathscr{K})$ des angles non aigus.*

Convenons d'appeler le cône $\mathscr{K}^*$ *conjugué* par rapport à $\mathscr{K}$.

La question se pose naturellement de savoir : qu'est-ce que représente donc le cône conjugué par rapport à $\mathcal{K}^*$ c'est-à-dire que peut-on dire de l'ensemble $(\mathcal{K}^*)^*$? De la définition du cône $\mathcal{K}^*$ il s'ensuit que l'ensemble $(\mathcal{K}^*)^*$ doit contenir l'ensemble initial $\mathcal{K}$ (pourquoi?). Mais il importe de savoir si ces deux ensembles coïncident ou non. De plus, en cherchant à vérifier ce fait à l'aide des raisonnements géométriques, on voit que ce n'est pas si simple. Quoi qu'il en soit, pour démontrer la coïncidence des ensembles $(\mathcal{K}^*)^*$ et $\mathcal{K}$ nous recourrons à la méthode algébrique, qui se fonde sur le théorème 1 ; au fond, comme on le verra ci-après, le *contenu géométrique du théorème* 1 *dans le cas d'un système à trois inconnues consiste précisément dans l'égalité* $(\mathcal{K}^*)^* = \mathcal{K}$.

Démontrons donc le théorème suivant.

● THÉORÈME 2. *Soit* $\mathcal{K}$ *un cône polyèdre convexe. L'ensemble* $(\mathcal{K}^*)^*$ *coïncide alors avec* $\mathcal{K}$.

Ce théorème peut être énoncé d'une autre façon. Désignons par $\mathcal{K}_1$ le cône $\mathcal{K}$ et par $\mathcal{K}_2$ celui $\mathcal{K}^*$. Ce théorème montre donc que

$$\text{si } \mathcal{K}_1^* = \mathcal{K}_2, \text{ alors } \mathcal{K}_2^* = \mathcal{K}_1,$$

ou : si un cône est conjugué par rapport à un autre, ce dernier l'est aussi par rapport au premier, en d'autres termes, la *relation de conjugaison est réciproque.*

● DÉMONSTRATION DU THÉORÈME 2. Soit $C\ (a,\ b,\ c)$ un point arbitraire de l'ensemble $(\mathcal{K}^*)^*$. Pour chaque point $B\ (x,\ y,\ z) \in \mathcal{K}^*$ doit se remplir l'inégalité $(B,\ C) \leqslant 0$, c'est-à-dire

$$ax + by + cz \leqslant 0. \qquad (11)$$

Mais le fait même que le point B appartient à l'ensemble $\mathcal{K}^*$ signifie, comme nous l'avons dit plus haut, l'observation des inégalités (10). Ainsi, chaque solution $(x,\ y,\ z)$ du systè-

me (10) doit vérifier l'inégalité (11) aussi. En d'autres termes, l'*inégalité* (11) *est une conséquence du système* (10).

D'après le théorème 1 démontré ci-dessus cela n'est possible que si l'inégalité (11) est une combinaison linéaire des inégalités (10), c'est-à-dire lorsque

$$(a,\ b,\ c) = t_1(a_1,\ b_1,\ c_1) + t_2(a_2,\ b_2,\ c_2) + \ldots \\ \ldots + t_m(a_m,\ b_m,\ c_m),$$

où $t_1, t_2, \ldots, t_m$ sont des nombres non négatifs quelconques. Mais cette dernière égalité signifie que

$$C = t_1A_1 + t_2A_2 + \ldots + t_mA_m,$$

c'est-à-dire le point C appartient au cône $\mathscr{K}$. Ainsi, tout point C appartenant à $(\mathscr{K}^*)^*$ appartient également à $\mathscr{K}$. Mais d'après ce qui a été exposé auparavant $\mathscr{K}$ appartient à $(\mathscr{K}^*)^*$. Par conséquent, $\mathscr{K} = (\mathscr{K}^*)^*$. Le théorème est ainsi démontré.

§ 10. THÉORÈME DE DUALITÉ DANS LA PROGRAMMATION LINÉAIRE

La programmation linéaire est un domaine relativement récent des mathématiques appliquées dont le développement a été imposé par la résolution de différents problèmes économiques.

Les problèmes qui se posent, en règle générale, dans l'économie et surtout dans la planification de la production sont surtout des problèmes aux extrêmaux consistant en la recherche de *solutions optimales.* En simplifiant (ou en exagérant même) la situation réelle, admettons que le rendement de l'atelier de montage d'une entreprise quelconque soit 100 articles du premier modèle ou 300 articles du second modèle par jour (l'entreprise est spécialisée dans la production de ces deux modèles), tandis que le service de contrôle ne peut examiner que 150 articles (de tout modèle) par jour. On sait

que les articles du premier modèle coûtent deux fois plus cher. Il faut donc planifier le travail de l'entreprise de façon à lui assurer les profits maximums. Tout récemment encore l'unique méthode de résolution de tels problèmes consistait à évaluer empiriquement, ou bien à traiter toutes les variantes possibles pour en choisir ensuite la meilleure. Maintenant la situation a radicalement changé. Ces dernières décennies toute production est devenue tellement compliquée qu'un simple traitement des variantes s'avère impossible, car le nombre de facteurs qui interviennent dans la résolution d'un tel problème se calcule aujourd'hui par des milliards. Ceci a conditionné l'intérêt accru porté aux méthodes mathématiques de résolution des problèmes économiques. L'apparition des calculateurs modernes et les progrès réalisés ces derniers temps par la technique de calcul ont grandement favorisé la « mathématisation de l'économie ».

Revenons tout de même à notre exemple. Le plan de production cherché est défini par deux nombres entiers non négatifs x, y (x étant le nombre d'articles du premier et y du deuxième modèle) qui doivent vérifier les conditions suivantes:

1) $3x + y \leqslant 300$ *);
2) $x + y \leqslant 150$;
3) $2x + y$ doit être maximal.

En d'autres termes, parmi les solutions entières non négatives du système

$$\left.\begin{aligned} 3x + y &\leqslant 300, \\ x + y &\leqslant 150 \end{aligned}\right\} \tag{1}$$

*) Cette condition est imposée par l'atelier de montage. En effet, au lieu d'un seul article du premier modèle cet atelier peut produire trois articles du deuxième. Par conséquent, toute la production de l'atelier en articles du deuxième type est $3x + y$ pièces; ce nombre ne doit dépasser 300.

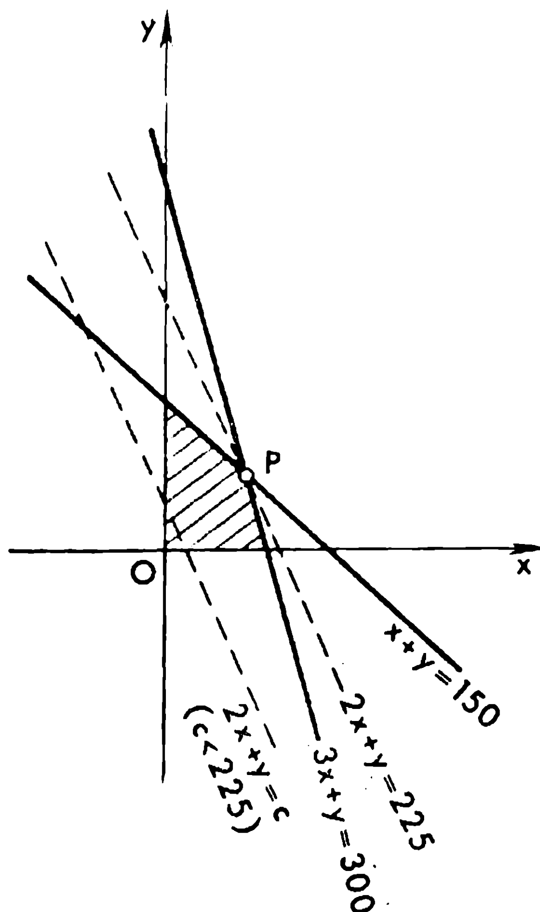


Fig. 50

il faut choisir celle qui maximise la fonction linéaire

$$f = 2x + y.$$

Si l'on adopte dans un plan un système de coordonnées rectangulaires xOy, l'ensemble des solutions du système (1) prendra alors la forme d'un polygone hachuré sur la figure 50. De ce même dessin il résulte que le problème considéré a le point P (75, 75), un des sommets du polygone, pour solution.

En effet, examinons la droite $2x + y = c$, c étant un nombre quelconque. Désignons-la par l_c. Avec l'augmentation du nombre c la droite l_c se déplace « vers le haut » (tout en restant parallèle à sa position initiale). La valeur maximale de c pour laquelle la droite l_c possède encore des points communs avec le polygone hachuré est celle

où cette droite passe par le point P. Par conséquent, en ce point la fonction $2x + y$ atteint sa valeur maximale (par rapport à ses valeurs en d'autres points du polygone).

L'exemple examiné est, certes, primitif mais il permet quand même de se faire une idée sur le caractère des problèmes de la programmation linéaire. Chacun d'eux consiste à maximiser (ou minimiser) la valeur d'une fonction linéaire à n variables

$$f = c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

à condition que ces variables obéissent à un certain système d'inégalités linéaires (qui doivent nécessairement comprendre les conditions des variables non négatives: $x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$).

Les méthodes de résolution des problèmes de la programmation linéaire sont actuellement élaborées en détail. La plupart de ces méthodes concordent bien avec le programme scolaire. Il s'agit ici d'un chapitre extrêmement important de l'algèbre élémentaire dont l'objet est d'étudier les *systèmes d'équations et d'inégalités linéaires.*

L'auteur ne se propose pas pour but d'exposer ici les fondements de la programmation linéaire, ceux-ci étant traités dans toute une série d'ouvrages spécialisés. Nous n'examinerons ici qu'un seul principe qui semble être du point de vue mathématique le plus important dans la programmation linéaire, c'est le *principe de dualité.*

Nous commençons par énoncer le problème de la programmation linéaire sous sa forme générale.

Soit un système

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + b_1 \geqslant 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + b_2 \geqslant 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + b_m \geqslant 0 \end{array}\right\} \qquad (2)$$

de m inégalités linéaires à n inconnues *) ainsi qu'une fonction linéaire

$$f = c_1x_1 + c_2x_2 + \ldots + c_nx_n.$$

De toutes les solutions non négatives ($x_1 \geqslant 0$, $x_2 \geqslant 0$, ..., $x_n \geqslant 0$) *du système* (2) *il faut choisir celle qui maximise la fonction* f.

Convenons d'appeler le problème posé *problème initial* ou *problème A*.

Associons le problème A avec un autre problème que nous appellerons *dual* par rapport à A ou *problème A'*. Voici donc sa formulation : soit un système

$$\left.\begin{array}{l} a_{11}x_1 + a_{21}y_2 + \ldots + a_{m1}y_m + c_1 \leqslant 0, \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m + c_2 \leqslant 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m + c_n \leqslant 0 \end{array}\right\} \quad (2')$$

de n inégalités linéaires à m inconnues et une fonction linéaire

$$\varphi = b_1y_1 + b_2y_2 + \ldots + b_my_m.$$

De toutes les solutions non négatives du système (2) *il faut choisir celle qui minimise la fonction* φ.

En comparant les problèmes A et A' nous voyons que :

1) Le coefficient de la j-ième inconnue dans la i-ème équation du système (2) coïncide avec celui de la i-ème inconnue dans la j-ième équation du système (2').

2) Les termes libres des inégalités de ces problèmes se confondent avec les coefficients des inconnues de la fonction linéaire de l'autre problème.

3) Dans le système d'inégalités du problème A, où l'on demande de maximiser f, toutes les inégalités sont du type

---

*) Dans le système (2) les désignations sont les suivantes : le coefficient de la j-ième inconnue dans la i-ème inégalité est noté a_{ij}.

$\geqslant 0$. Dans celui du problème A', où l'on demande de minimiser φ, toutes les inégalités sont, par contre, du type $\leqslant 0$.

L'un des théorèmes fondamentaux de la programmation linéaire, *théorème de dualité*, s'énonce ainsi.

● THÉORÈME DE DUALITÉ. *Si le problème initial possède une solution, le problème dual a aussi une solution. Le maximum de la fonction f est égal au minimum de la fonction φ:*

$$\max f = \min \varphi.$$

Nous allons démontrer ce théorème en le réduisant au problème de la compatibilité d'un système d'inégalités.

Pour simplifier la démonstration, nous la ferons par étapes.

● ÉTAPE 1. LEMME. *Si $x_1^0, \ldots, x_n^0$ est une solution non négative du système* (2) *et $y_1^0, \ldots, y_m^0$ une solution non négative du système* (2′), *les valeurs des fonctions f et φ pour ces solutions sont liées par une inégalité*

$$f_0 \leqslant \varphi_0.$$

● DÉMONSTRATION. Considérons les inégalités du système (2) où les valeurs $x_1, \ldots, x_n$ sont remplacées par $x_1^0, \ldots, x_n^0$. Multiplions la première de ces inégalités par y_1^0, la deuxième par y_2^0, etc., et additionnons ensuite toutes les inégalités ainsi obtenues:

$$(a_{11}y_1^0x_1^0 + \ldots + a_{mn}y_m^0x_n^0) + b_1y_1^0 + \ldots + b_mx_m^0 \geqslant 0$$

(il faut tenir compte de ce que nous multiplions les inégalités par des nombres non négatifs, les signes d'inégalités restent donc inchangés). De façon analogue, multiplions la première inégalité du système (2′) par x_1^0, la deuxième — par x_2^0, etc., et additionnons-les ensuite:

$$(a_{11}y_1^0x_1^0 + \ldots + a_{mn}y_m^0x_n^0) + c_1x_1^0 + \ldots + c_nx_n^0 \leqslant 0.$$

Dans les deux cas l'expression figurant entre parenthèses est égale à la somme des termes de la forme $a_{ij}y_i^0x_j^0$ pour tous les $i = 1, \ldots, m$; $j = 1, \ldots, n$. Par conséquent, les deux expressions entre parenthèses coïncident. Mais alors on a

$$c_1x_1^0 + \ldots + c_nx_n^0 \leqslant b_1y_1^0 + \ldots + b_my_m^0$$

ou $f_0 \leqslant \varphi_0$. Le lemme est démontré.

● ÉTAPE 2. *La réduction des problèmes A et A' à la résolution d'un système d'inégalités.*

Examinons le système d'inégalités « combiné » suivant:

$$\left.\begin{array}{l|l}
a_{11}x_1 + \ldots + a_{1n}x_n & \qquad\qquad + b_1 \geqslant 0, \\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\
a_{m1}x_1 + \ldots + a_{mn}x_n & \qquad\qquad + b_m \geqslant 0, \\
\hline
 & a_{11}y_1 + \ldots + a_{m+1}y_m + c_1 \leqslant 0 \\
 & \ldots\ldots\ldots\ldots \\
 & a_{1n}y_1 + \ldots + a_{mn}y_m + c_n \leqslant 0, \\
\hline
c_1x_1 + \ldots + c_nx_n & -b_1y_1 - \ldots - b_my_m \qquad \geqslant 0.
\end{array}\right\} \qquad (S)$$

Il se compose des systèmes (2), (2') et de l'inégalité $f - \varphi \geqslant 0$. Les inconnues du système (S) sont $x_1, \ldots, x_n$; $y_1, \ldots, y_m$ ($n + m$ inconnues en tout). Nous allons maintenant établir le fait suivant.

Si le système (S) a une solution non négative $x_1^0, \ldots, x_n^0$, $y_1^0, \ldots, y_m^0$, les nombres $x_1^0, \ldots, x_n^0$ donnent la solution du problème A, $y_1^0, \ldots, y_m^0$ celle du problème A', $f_0 = \varphi_0$.

Arrêtons-nous sur cet endroit pour souligner l'importance de cette proposition. Ce qui importe dans celle-ci c'est que le *problème de la programmation linéaire, c'est-à-dire celui de maximisation, se réduit à la résolution d'un système d'inégalités linéaires sans aucune exigence, quelle qu'elle soit, de maximisation.* En réalité, la résolution du système (S) (dans le domaine des valeurs non négatives des inconnues)

n'est certainement en rien plus facile que celle du problème initial de la programmation linéaire (problème A) ; cependant, le fait même de la possibilité d'une telle réduction est d'un intérêt particulier.

Démontrons donc la proposition avancée ci-dessus. Il est évident tout d'abord que les nombres $x_1^0, \ldots, x_n^0$ sont non négatifs et vérifient le système (2) ; de manière analogue, les nombres $y_1^0, \ldots, y_m^0$ le sont également et satisfont à (2'). D'autre part ces nombres vérifient l'inégalité

$$f_0 \geqslant \varphi_0$$

(qui découle de la dernière inégalité du système (S)). D'autre part, en vertu du lemme on a

$$f_0 \leqslant \varphi_0.$$

Par conséquent, $f_0 = \varphi_0$.

Ensuite, si $x_1, \ldots, x_n$ est une solution non négative *quelconque* du système (2), on obtient alors d'après le lemme

$$f \leqslant \varphi_0.$$

En comparant ce résultat avec $f_0 = \varphi_0$, on a $f \leqslant f_0$ d'où il résulte que f_0 est la valeur maximale de f.

De manière analogue, si $y_1, \ldots, y_m$ est une solution non négative *quelconque* du système (2'), on obtient donc en vertu de ce même lemme

$$f_0 \leqslant \varphi ;$$

en confrontant ceci avec $f_0 = \varphi_0$, on a $\varphi_0 \leqslant \varphi$, c'est-à-dire φ_0 est le minimum de φ. Ceci démontre la proposition énoncée ci-dessus.

● ÉTAPE 3. *L'achèvement de la démonstration.* Il ne nous reste qu'à démontrer le fait suivant : si le problème A a une solution, le système (S) a une solution non négative. Car dans ce cas, comme nous l'avons montré plus haut, $f_0 = \varphi_0$, c'est-à-dire $\max f = \min \varphi$.

Raisonnons par l'absurde, c'est-à-dire supposons que le

système (S) ne possède pas de solutions non négatives. Mais pour ce cas il existe le corollaire du théorème sur les systèmes incompatibles (§ 8). Il est vrai que celui-ci se rapporte à un système dont les inégalités sont du type $\geqslant 0$ et dans notre système (S) il y a des inégalités du type $\leqslant 0$. Mais pour parer à cet inconvénient il faut écrire le système (S) sous la forme

$$\left.\begin{array}{ll}
a_{11}x_1 + \ldots + a_{1n}x_n & + b_1 \geqslant 0, \\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\
a_{m1}x_1 + \ldots + a_{mn}x_n & + b_m \geqslant 0, \\
 & -a_{11}y_1 - \ldots - a_{m1}y_m - c_1 \geqslant 0, \\
 & \ldots\ldots\ldots\ldots \\
 & -a_{1n}y_1 - \ldots - a_{mn}y_m - c_n \geqslant 0, \\
c_1x_1 + \ldots + c_nx_n & -b_1y_1 - \ldots - b_my_m \geqslant 0.
\end{array}\right\} \quad (S')$$

Ainsi, supposons que le système (S') ne possède pas de solutions non négatives. D'après le corollaire du théorème sur les systèmes incompatibles, il existe des nombres non négatifs $k_1, \ldots, k_m, l_1, \ldots, l_n, s$ $(m + n + 1$ nombres en tout) tels que *)

$$\left.\begin{array}{l}
a_{11}k_1 + \ldots + a_{m1}k_m + c_1s \leqslant 0, \\
\ldots\ldots\ldots\ldots \\
a_{1n}k_1 + \ldots + a_{mn}k_m + c_ns \leqslant 0;
\end{array}\right\} \quad (3)$$

$$\left.\begin{array}{l}
-a_{11}l_1 - \ldots - a_{1n}l_n - b_1s \leqslant 0, \\
\ldots\ldots\ldots\ldots \\
-a_{m1}l_1 - \ldots - a_{mn}l_n - b_ms \leqslant 0;
\end{array}\right\} \quad (3')$$

$$b_1k_1 + \ldots + b_mk_m - c_1l_1 - \ldots - c_nl_n < 0. \quad (4)$$

---

*) Les nombres $k_1, \ldots, k_m, l_1, \ldots, l_n, s$ sont précisément ceux par lesquels nous multiplions les première, deuxième, ..., $(m + n + 1)$-ième inégalités du système (S') pour obtenir (après l'addition) l'inégalité contradictoire $a_1x_1 + \ldots + a_nx_n + a'_1y_1 + \ldots + a'_my_m + d \geqslant 0$, où $a_1, \ldots, a_n, a'_1, \ldots, a'_m \leqslant 0$, d étant un nombre négatif.

Montrons tout d'abord que le nombre s n'est pas nul. En effet, admettons qu'il n'en est pas ainsi, c'est-à-dire que $s = 0$. Considérons une solution non négative quelconque $x_1^0, \ldots, x_n^0$ du système (2) et une solution non négative quelconque $y_1^0, \ldots, y_m^0$ du système (2'). En raisonnant tout comme lors de la démonstration du lemme, on obtient que

$$(a_{11}k_1x_1^0 + \ldots + a_{mn}k_mx_n^0) + b_1k_1 + \ldots + b_mk_m \geqslant 0.$$

Mais l'expression figurant entre parenthèses est $\leqslant 0$ (on l'obtient si l'on multiplie la première des inégalités (3) par x_1^0, la deuxième par x_2^0, etc., et on additionne ensuite les inégalités obtenues). D'où

$$b_1k_1 + \ldots + b_mk_m \geqslant 0. \tag{5}$$

De façon analogue on trouve que

$$(a_{11}y_1^0l_1 + \ldots + a_my_m^0l_n) + c_1l_1 + \ldots + c_nl_n \leqslant 0,$$

et comme l'expression entre parenthèses $\geqslant 0$ (on l'obtient en multipliant la première des inégalités (3') par y_1^0, la deuxième par y_2^0, etc., et en les additionnant ensuite), on a

$$c_1l_1 + \ldots + c_nl_n \leqslant 0. \tag{6}$$

Mais les inégalités (5) et (6) se trouvent en contradiction avec (4).

Ainsi, s n'est pas nul. Dans ce cas, il s'ensuit de (3) que les nombres $\frac{k_1}{s}, \ldots, \frac{k_m}{s}$ fournissent une solution non négative du système (2), et de (3') que les nombres $\frac{l_1}{s}, \ldots, \frac{l_n}{s}$ représentent une solution non négative du système (2') et, enfin, de (4) que pour ces solutions $\varphi - f < 0$. Mais cela contredit le lemme. Ainsi, en supposant que le système (S) ne possède pas de solutions non négatives, on est amené à une contradiction. Par conséquent, de telles solutions existent nécessairement, et le théorème de dualité est ainsi démontré.

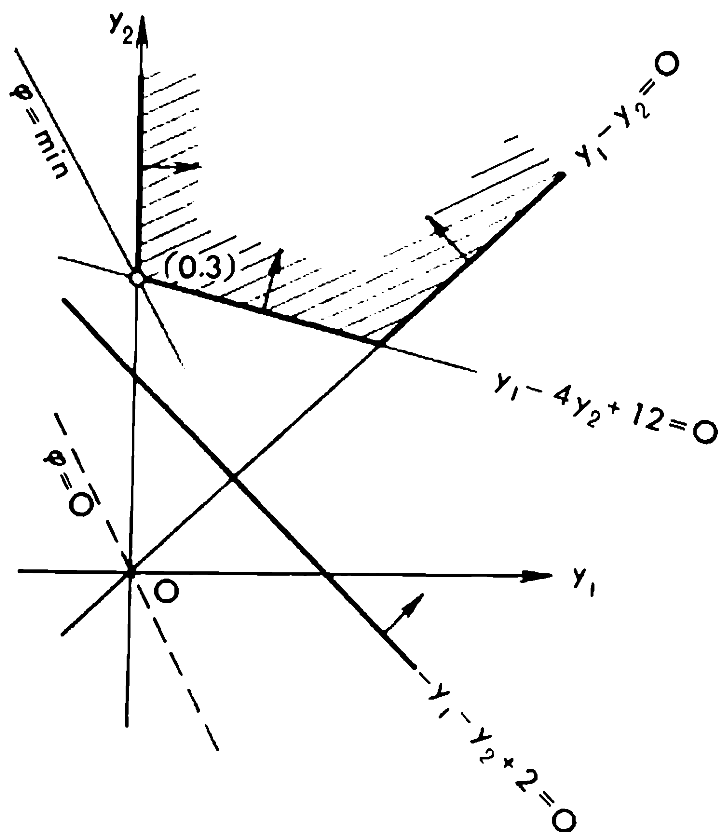


Fig. 51

● EXEMPLE. Trouver le maximum de la fonction

$$f = 2x_2 + 12x_3$$

sous l'hypothèse que les variables x_1, x_2, x_3 sont non négatives et satisfont aux inégalités

$$\left.\begin{aligned} x_1 - x_2 - x_3 + 2 \geqslant 0, \\ -x_1 - x_2 - 4x_3 + 1 \geqslant 0. \end{aligned}\right\}$$

● SOLUTION. Convenons d'appeler le problème posé problème A. Son dual (A') doit s'énoncer ainsi : trouver le minimum de la fonction

$$\varphi = 2y_1 + y_2$$

sous l'hypothèse que les variables y_1, y_2 sont non négatives et satisfont aux inégalités

$$\left.\begin{aligned} y_1 - y_2 \qquad &\leqslant 0, \\ -y_1 - y_2 + 2 \quad &\leqslant 0, \\ -y_1 - 4y_2 + 12 &\leqslant 0. \end{aligned}\right\} \qquad (7)$$

Le problème A' peut se résoudre graphiquement en représentant sur le plan des coordonnées y_1Oy_2 le domaine des solutions du système (7) (fig. 51). Sur ce même dessin on voit que la fonction φ possède son minimum au point (0, 3), un des sommets de ce domaine. Cette valeur est égale à —3. D'après le théorème de dualité, le maximum de la fonction f doit aussi être égal à —3.

TABLE DES MATIÈRES

A NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur le contenu de ce livre, sa traduction et sa présentation, ainsi que toute autre suggestion. Notre adresse : Editions Mir, 2, Pervi Rijski péréoulok, Moscou, I-110, GSP, U.R.S.S.

Imprimé en Union Soviétique

